# Splicing of Effector Cell Protease Receptor-1 mRNA Is Modulated by an Unusual Retained Intron[†,‡]

Dario C. Altieri[*]

*Department of Vascular Biology, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037, and The Boyer Center for Molecular Medicine, Department of Pathology, Yale University School of Medicine, 295 Congress Avenue, New Haven, Connecticut 06536*

*Received July 6, 1994; Revised Manuscript Received September 13, 1994*[⊗]

ABSTRACT: Effector cell protease receptor-1 (EPR-1) is a transmembrane glycoprotein receptor for factor Xa that contributes to cell surface assembly of proteolytic activities and leukocyte mitogenesis. It is now shown that membrane expression of EPR-1 is dynamically modulated by mRNA splicing. Northern hybridization analysis of EPR-1-expressing cells and genetically engineered transfectants demonstrates that this mechanism involves removal of a 451 bp intervening sequence retained in 70−90% of mature mRNA, as quantitated by polymerase chain reaction amplification and ribonuclease protection studies. Splicing of the intervening sequence occurs in a cell type-specific fashion, as judged by the constitutive membrane overexpression of EPR-1 in certain leukemic B lymphocytes and monocytic cells. Furthermore, phenotypic analysis of cell lines stably transfected with functionally spliced or unspliced EPR-1 constructs suggests a potential role of intron *cis*-acting sequence(s) in splicing regulation. Instead of a transmembrane receptor for factor Xa (EPR-1a), the most prevalent unspliced EPR-1 transcript generates a novel truncated protein of 110 amino acids (EPR-1b), in which a unique intron-encoded -COOH terminus carries a potential nuclear targeting signal PPQHRAKS. An antibody generated against the intron-encoded sequence of EPR-1b demonstrates prominent nuclear localization of this variant isoform in indirect immunofluorescence staining of permeabilized cells. These findings provide evidence for a novel mechanism based on high efficiency intron retention modulating factor Xa-dependent cellular effector functions.

The assembly of proteolytic activities on vascular cells preserves the homeostasis of internal body fluids (Furie & Furie, 1988) and participates in a variety of cellular effector functions. Binding of coagulation and fibrinolytic proteases to cell surface protease receptors transduces intracellular signals (Vu et al., 1991), initiates transcription of early activation-dependent genes (Daniel et al., 1986), influences normal and malignant cell motility (Bar-Shavit et al., 1983; Ossowski, 1987), and stimulates DNA synthesis and cell proliferation (Glenn & Cunningham, 1979; Kirchheimer et al., 1989). A recently identified molecule, designated effector cell protease receptor-1 (EPR-1),[1] mediates the interaction of the blood protease factor Xa with monocytes and various leukocyte subpopulations, thus promoting membrane assembly of proteolytic activities (Altieri & Edgington, 1990; Worfolk et al., 1992). Furthermore, consistent with the mitogenic effect of factor Xa on mesenchymal cells (Gasic et al., 1992), occupancy of EPR-1 with factor Xa or with selected monoclonal antibodies (mAbs) costimulates lymphocyte proliferation in vitro (Altieri & Stamnes, 1994). Using a functional cloning strategy of the cDNA, the primary structure of EPR-1 has been recently elucidated and defines a novel transmembrane glycoprotein receptor for factor Xa of 337 amino acids, potentially implicated in protease-dependent intracellular signaling mechanisms (Altieri, 1994). It is now shown that EPR-1 expression is modulated by an unusual pathway of mRNA splicing (Sharp, 1987; Green, 1991; Maniatis, 1991), which involves an intervening sequence almost invariably retained in mature mRNA. This mechanism introduces a profound protein isoform diversity. The most prevalent and ubiquitous unspliced EPR-1 transcript generates a novel, nuclear-localized, truncated protein with a unique intron-encoded -COOH terminus.

## EXPERIMENTAL PROCEDURES

*Cells and Cell Cultures.* Peripheral blood mononuclear cells (PBMC) were isolated from heparinized blood drawn from informed normal healthy volunteers or from a patient (M.W.) with chronic lymphocytic leukemia (CLL) by differential centrifugation on Ficoll-Hypaque gradient (Sigma Chemical Co., St. Louis, MO), as described (Altieri & Edgington, 1990). PBMC at $5 \times 10^6$/mL were activated by a 2−7 day culture in 24-well tissue culture plates (Costar Corp., Cambridge, MA) in the presence of 20 μg/mL phytohemagglutinin (PHA, Calbiochem Corp., La Jolla, CA). Monocytic THP-1 cells, erythroleukemia HEL, B lymphoma Daudi, T leukemia MLT (American Type Culture Collection, ATCC, Rockville, MD) were grown in RPMI 1640 tissue culture medium (Whittaker Bioproducts, Walkersville, MD)

[†] This work was supported by NIH Grants RO1 HL-43773 and HL-54131. This work was done during the tenure of Established Investigatorship award from the American Heart Association to Dr. Altieri.

[‡] The sequence data contained in this paper has been deposited to GenBank with accession number L32866.

* Address correspondence to Dario C. Altieri, M.D., The Boyer Center for Molecular Medicine, Department of Pathology, Yale University School of Medicine, 295 Congress Ave., New Haven, Connecticut 06536. Telephone: (203)737-2869. Fax: (203)737-2290.

[⊗] Abstract published in *Advance ACS Abstracts,* October 15, 1994.

[1] Abbreviations: CHO, chinese hamster ovary; CLL, chronic lymphocytic leukemia; EPR-1, effector cell protease receptor-1; HUVEC, human umbilical vein endothelial cells; mAb, monoclonal antibody; ORF, open reading frame; PBMC, peripheral blood mononuclear cells; PCR, polymerase chain reaction; PHA, phytohemagglutinin; WT, wild type.

0006-2960/94/0433-13848$04.50/0 

plus 1 mM L-glutamine (Whittaker), 10% fetal calf serum (FCS, Whittaker), and $10^{-5}$ M $\beta$-mercaptoethanol (Eastman Kodak Co., Rochester, NY). Monocytic WEHI.3 cells (ATCC) were grown in DMEM medium (Whittaker) plus 10% FCS, 1 mM L-glutamine, nonessential amino acids, and penicillin-streptomycin (Irvine Scientific, Santa Ana, CA). Cell surface expression of EPR-1 was quantitated by flow cytometry using anti-EPR-1 mAb 2E1 (Altieri, 1994).

*Molecular Cloning of the cDNA for EPR-1.* The experimental strategy to isolate cDNA clones encoding EPR-1 has been described previously (Altieri, 1994). Thirty-two independent clones were isolated from four different oligo-dT-primed human cDNA libraries, and all sequenced on both strands of exonuclease III-generated (Promega Corp., Madison, WI) nested deletions using Sequenase (United States Biochemical, Cleveland, OH).

*Transfection Experiments.* EPR-1 cDNA constructs in the eukaryotic expression vector pRC/CMV (Invitrogen), were transfected (15 $\mu$g of plasmid DNA) in subconfluent cultures of Chinese Hamster Ovary (CHO) cells by electroporation. Forty-eight hours after transfection, cells were harvested, selected in DMEM plus 1 mg/mL Geneticin (G418, GIBCO, Gaithersburg, MD) at $6 \times 10^3$cells/80 mm diameter tissue culture Petri dishes, and cloned by limiting dilution in 96-well microtiter plates (Costar) at 0.5−3 cells/well in selection medium.

*Polymerase Chain Reaction (PCR).* The following EPR-1-derived oligonucleotides were used: forward A10/25, 5′(-11)GAACAGCCGAGATGACCTCC(9)3′; forward C3/27, (81)5′ TCATCTTACGCCAGACTTCAGCC (103)3′; reverse G3/27, (784)5′ CGATCTCGACTGTCGAAACAAGC (762)3′; forward T5/27, (161)5′ CGTGTGGAGAACGTGACAGATGTG (184)3′; reverse G5/27, (739)5′ AATGTGTCTGGACCTCATGTTG (718)3′; junctional GA3 reverse (696)5′ AACCGTCCACG|ACCGACCCGA (217)3′ *Sma*I reverse (839)5′ GATTCCCGGGCTTACCAGGTGAGAAG (814)3′ (positions in the EPR-1 mRNA are in parentheses). First strand cDNA was synthesized from total or poly(A)$^+$ mRNA isolated from EPR-1-expressing cells. Typically, 1 $\mu$g of poly(A)$^+$ mRNA or 16 $\mu$g of total RNA were incubated with 10 mM MeHgOH for 10 min at 22 °C, neutralized in 700 mM $\beta$-mercaptoethanol, and annealed with various EPR-1-derived oligonucleotides for 2 min at 65 °C. The reverse transcription reaction was carried out in a volume of 20 $\mu$L in the presence of 10 mM Tris-HCl, 50 mM KCl, pH 8.8, 0.1% Triton X-100, 5 mM $MgCl_2$, 1.25 mM dNTPs, 1 unit of RNAse inhibitor (RNAsin, Promega Corp., Madison, WI), and 15 units/reaction AMV reverse transcriptase (Stratagene, San Diego, CA) for 1 h at 42 °C. At the end of the incubation, the DNA−RNA hybrids were denatured by heating at 95 °C for 3 min and subjected to a second round of cDNA synthesis. Twenty or fifty percent of the synthesized cDNA was amplified by PCR using 0.5 $\mu$g of EPR-1-specific oligonucleotides in the presence of 10 mM KCl, 20 mM Tris-HCl, pH 8.8, 10 mM $(NH_4)SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100, 100 $\mu$g/mL nonacetylated bovine serum albumin, 400 $\mu$M dNTPs, and 2 units/reaction of $Vent_R$ DNA polymerase (New England Biolabs, Beverly, MA). Thirty-five cycles of amplification were carried out in a Perkin-Elmer Cetus 480 thermocycler with denaturation at 94 °C for 1 min, annealing at 54−66 °C for 1 min, and extension at 72 °C for 1 min. In separate experiments, aliquots of human cDNA libraries ($\lambda$gt11 MLT, HEL or human umbilical vein endothelial cells, HUVEC) or from a $\lambda$FIX II genomic library (human placenta, Stratagene) were used as templates for amplification at 5 $\mu$L/reaction. All PCR products generated in this study were analyzed on 1−2% analytical agarose gels (GIBCO, Gaithersburg, MD) by ethidium bromide staining, cloned in pCRII (Invitrogen) or pCRScript (Stratagene) and sequenced on both strands using M13- or EPR-1-specific primers. In some experiments, PCR products were transferred to GeneScreen (New England Nuclear, Boston, MA) for Southern hybridization. Nucleic acids were UV cross-linked (Stratalinker, Stratagene), prehybridized in 5× SSC, 5× Denhardt's solution, 0.1% sodium pyrophosphate, 0.5% SDS, and 100 $\mu$g/mL denatured salmon sperm DNA (Promega) for 2 h at 65 °C and hybridized with various $^{32}$P-random primed (Klenow, Boehringer-Mannheim, Indianapolis, IN)-labeled EPR-1 probes or internal EPR-1 oligonucleotides end-labeled with T4 polynucleotide kinase (Promega) for 6−12 h at 65 °C in a rolling hybridization oven (Hoefer Scientific, San Francisco, CA). The transfer membrane was washed twice in 5% SDS, 2× SSC at 65 or 68 °C for 30 min and once in 0.2× SSC at 22 °C, before exposure for autoradiography using a Kodak X-Omat AR X-ray film (Eastman Kodak) and intensifying screens.

*RNA Isolation, Northern Blots, and Ribonuclease Protection Assay.* Total RNA was extracted from EPR-1-expressing cells by the guanidinium-isothiocyanate method. Poly(A)$^+$ mRNA was prepared by magnetic separation on oligo-dT-coupled beads (Promega). RNA samples (0.5−1 $\mu$g of mRNA or 8−16 $\mu$g of total RNA) were electrophoresed on agarose formaldehyde gels in 20 mM $NaPO_4$ and 1 mM EDTA at 80 V for 4−5 h, transferred to GeneScreen, prehybridized, and hybridized with $^{32}$P-labeled EPR-1 probes as described above. For ribonuclease protection experiments, the pCRII plasmid containing the intronless EPR-1 probe PCR124 (see below) was linearized with *Xba*I upstream the 5′ end of the insert, and the antisense strand was transcribed in a total volume of 20 $\mu$L containing 40 mM Tris-HCl, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM DTT, 1 unit/reaction RNAse inhibitor (RNAsin, Promega), 10 mM ATP/GTP/CTP, 50−90 $\mu$Ci of [$^{32}$P]UTP (>3000 Ci/mmol, Amersham), and 10 units/reaction SP6 RNA polymerase for 1 h at 37 °C. After incubation with 10 units of DNase for 15 min at 37 °C, the PCR124 RNA transcript was separated on a 5% acrylamide/8 M urea denaturing gel, excised, and eluted in 1 mM EDTA, 0.5 M $NH_4OAc$, and 0.1% SDS overnight at 37 °C. Forty-thousand or eighty-thousand cpm of the PCR124 RNA transcript were hybridized with 8−24 $\mu$g of total RNA from HEL or THP-1 cells in 80% formamide, 100 mM sodium citrate, 300 mM sodium acetate, pH 6.4, and 1 mM EDTA for 8−14 h at 45 °C. At the end of the incubation, samples were treated with a 1:100 dilution of RNAse A/RNase T1 mixture (typically, 2 $\mu$g/mL RNAse T1, 40 $\mu$g/mL RNAse A) in 10 mM Tris-HCl, 5 mM EDTA, 300 mM NaCl (Ambion, Inc., Austin, TX) for 30 min at 37 °C. After precipitation, protected fragments were separated by electrophoresis on a 5% acrylamide/8 M urea denaturing gels, and visualized by autoradiography.

*Construction of a Functionally Spliced EPR-1 Minigene.* Ten micrograms of HEL total RNA were reverse transcribed with EPR-1-specific oligonucleotides and amplified in separate reactions with complementary splice-site specific oligonucleotides GA2 (forward) plus reverse *Sma*I oligonucleotide, or GA3 (reverse) plus forward A10/25 oligo-

nucleotide. The two resulting intronless PCR products of 248 bp and 172 bp, respectively, were fused together for 2 min at 65 °C and reamplified with forward oligonucleotide A10/25 plus reverse *Sma*I oligonucleotide. This construct of 420 bp comprised the 5′ translational initiation codon contained in the A10/25 oligonucleotide, and extended through the *Sma*I site at position 379 in the functionally spliced EPR-1 transcript (Figure 1A). The construct was cloned in the *Eco*RI and *Sma*I sites of pBSKS$^-$ and completely sequenced on both strands. A full length EPR-1 cDNA clone isolated by high stringency hybridization screening from a λgt11 MLT library (λ112) was digested with *Sac*I and *Sma*I and directionally cloned in-frame with the *Sma*I site of the intronless PCR construct. This functionally spliced EPR-1 minigene (F1, 1.1 kb) was sequenced on both strands and inserted in pRC/CMV.

*Synthetic Peptide and Sequence-Specific Antiserum.* A synthetic peptide designated P-25-A and duplicating the intron-encoded sequence Pro-Pro-Gln-His-Arg-Ala-Lys-Ser-Phe-Lys-Lys-Asp-Leu-Gly-His-Cys-Leu-Phe-Leu-Arg-Glu-Glu-Arg-Pro-Ala was synthesized and purified by high pressure liquid chromatography on C18 columns. Twenty-five milligrams of the P-25-A peptide were coupled in a 1:1 molar ratio to keyhole limpet hemocyanin (Calbiochem), and 500 μg of the conjugate in 1 mL was injected subcutaneously in multiple sites into a rabbit in complete Freund's adjuvant. After a 4-week interval, the animal was boosted with 500 μg of the conjugate subcutaneously in incomplete Freund's adjuvant and bled 1 week after the boost. Subsequent boosts were also followed by a 1-week interval before bleeding.

*Indirect Immunofluorescence.* HUVEC were grown to confluency onto 0.2% gelatin-coated glass coverslips for 4−5 days at 37 °C, washed in PBS, pH 7.2, and fixed with 1% paraformaldehyde for 30 min at 22 °C. After washes, cells were permeabilized with 0.5% Triton X-100 for 10 min at 22 °C, washed in PBS, pH 7.2, and incubated with 1:100 dilutions of control preimmune rabbit serum, anti-P-25-A sequence-specific antibody, or anti-Sm nuclear protein antiserum (generously provided by R. Ochs, The Scripps Research Institute) for 1 h at 22 °C. After washes, binding of the primary antibody was revealed by addition of a 1:20 dilution of rhodamine-conjugated goat anti-rabbit F(ab′)$_2$ fragments (Tago Inc., Burlingame, CA) for 1 h at 22 °C. Glass coverslips were mounted in SlowFade (Molecular Probes, Eugene, OR) and analyzed on a Leitz Diaplan fluorescence microscope.

## RESULTS

*The EPR-1 mRNA Contains a 451 bp Intervening Sequence.* A consensus DNA sequence deduced from 32 independent EPR-1 cDNA clones (Altieri, 1994) contained a single 5′ open reading frame (ORF) beginning with a translational initiation codon (Kozak, 1984) and extending for 330 bp before interruption by a ochre (UAA) termination codon (Figure 1A). Four additional in-frame termination codons were contained in a 451 bp intervening sequence present in all 32 EPR-1 cDNA clones (Figure 1A). Thirty-two bp downstream the last termination codon, a single ORF continued uninterrupted for 785 bp (Figure 1A), encoding the -COOH terminal two-thirds of the EPR-1 extracellular domain, a single membrane-spanning domain of 26 hydrophobic amino acids, and a serine-rich (26%), 81 amino acid-long cytoplasmic tail (Figure 1A; Altieri, 1994). No other forward or reverse ORF were identifiable in the EPR-1 cDNA, or in the intervening sequence. DNA sequencing of this intervening region revealed several clusters of adenosine residues (Figure 1B) but no significant homologies in available databases. To independently validate the authenticity of the termination codons and to rule out potential sequencing errors, a 454 bp *Pst*I subclone containing the EPR-1 intervening sequence (λ105), or a 850 bp *Pst*I fragment beginning downstream the last termination codon (λ114) were expressed as fusion proteins in *Escherichia coli*. While no recombinant molecule was generated in λ105, a $M_r$ ~31 kDa fusion protein was specifically induced in λ114 cultures, under the same experimental conditions (not shown).

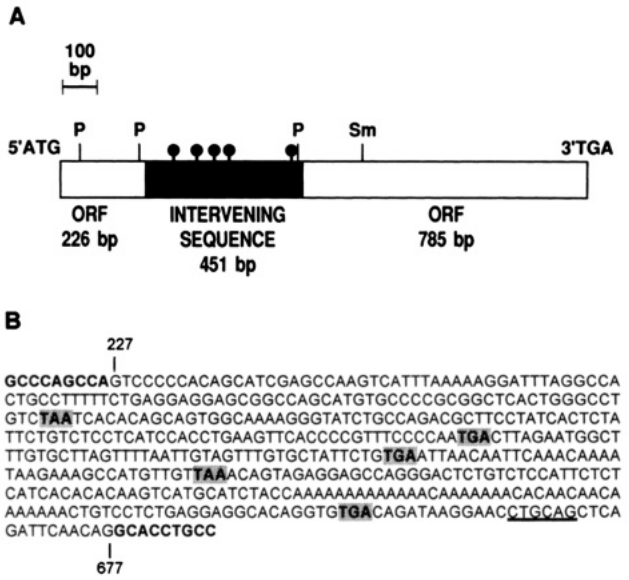


FIGURE 1: (A) Organization of the EPR-1 mRNA. The EPR-1 consensus DNA sequence was deduced from 32 independent clones isolated from four oligo dT-primed human cDNA libraries. A 5′ ORF beginning with a translational initiation codon (Kozak, 1984) is interrupted by five in-frame termination codons (●) contained in a 451 bp intervening sequence. Thirty-two bp downstream the last termination codon, the EPR-1 ORF continues uninterrupted for 785 bp. P, *Pst*I; S, *Sma*I. (B) cDNA sequence of the EPR-1 intervening sequence. In-frame termination codons are boxed. Translated sequences at the 5′ and 3′ junctions with the EPR-1 intervening sequence are in bold. A *Pst*I site is double underlined.

To determine the cellular distribution of this intervening sequence, oligo-dT-primed cDNA libraries or reverse-transcribed RNA of EPR-1$^+$ cells were subjected to PCR amplification. As shown in Figure 2, the oligonucleotide pair T+G5/27, flanking the EPR-1 intervening sequence, generated a single DNA product of 579 bp from all the templates analyzed, consistent with the size of the EPR-1 mRNA containing the intervening sequence (Figure 2). No additional lower molecular weight bands were detected by ethidium bromide staining (Figure 2, and see below).

*Characterization of the EPR-1 Intervening Sequence as a Retained Intron.* Despite the presence of the termination codons (Figure 1), transfection of CHO cells with the EPR-1 cDNA containing the intervening sequence (λ407) was specifically associated with surface expression of a recombinant molecule recognized by anti-EPR-1 mAbs (Altieri, 1994). Functionally, EPR-1 transfectants, but not wild type (WT) CHO cells, bound the natural EPR-1 ligand $^{125}$I-factor Xa (Altieri & Edgington, 1990; Worfolk et al., 1992) in a

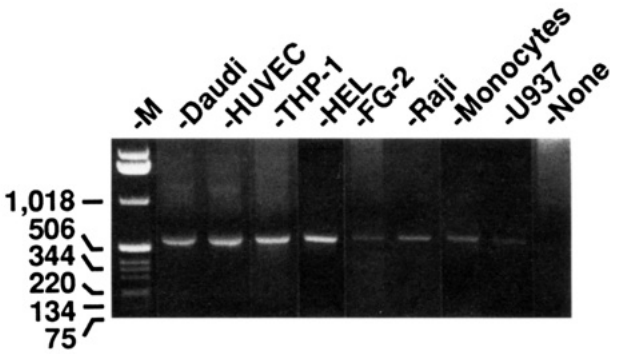


FIGURE 2: Cellular distribution of the EPR-1 intervening sequence. PCR amplification of EPR-1 sequences was carried out with the flanking oligonucleotide pair T+G5/27 from oligo-dT-primed cDNA libraries (FG-2, Raji, monocytes, U937) or reverse-transcribed RNA of EPR-1$^+$ cells (THP-1, Daudi, HEL, HUVEC). A single DNA product of 579 bp containing the EPR-1 intervening sequence is amplified from all templates analyzed, while no additional lower molecular weight DNA products were detected by ethidium bromid staining. None, no template control. M, molecular weight markers in bp.

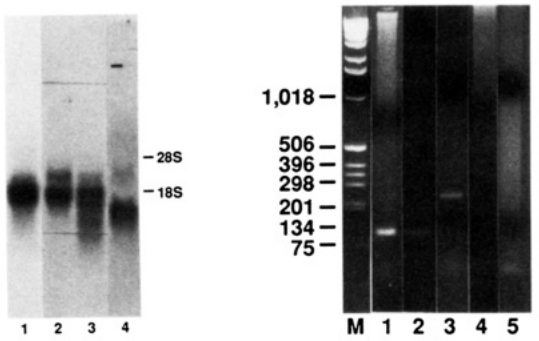


FIGURE 3: Functional processing of EPR-1 mRNA. (Left panel) Northern hybridization. Samples from EPR-1$^+$ HEL (1 $\mu$g of mRNA, lane 1), 2-day PHA-activated lymphocytes (16 $\mu$g of total RNA, lane 2), 3-day PHA-activated lymphocytes (16 $\mu$g of total RNA, lane 3), or EPR-1 CHO transfectants (16 $\mu$g of total RNA, lane 4) were electrophoresed on agarose-formaldehyde gels, transferred to GeneScreen, UV cross-linked, hybridized with a 1.6 kb EPR-1 probe, and exposed for autoradiography. The EPR-1 message in CHO transfectants appeared as a smaller transcript of 1.5 kb (lane 4), as opposed to the 1.9 kb band detected in other EPR-1-expressing cells (lanes 1−3). Position of ribosomal RNA is indicated on the right. (Right panel) Cell type-specific removal of the EPR-1 intervening sequence. Total RNA isolated from EPR-1 CHO transfectants, CLL lymphocytes, or WEHI.3 monocytes, was reverse-transcribed, primed with EPR-1-specific oligonucleotides, and amplified with the flanking oligonucleotide pairs G+T5/27 (lanes 1 and 2), or C+G3/27 (lane 3). Single DNA products of 128 and 253 bp were amplified from CHO transfectants (lane 1), CLL lymphocytes (lane 2), or from WEHI.3 monocytes (lane 3), respectively. Control reactions in which no template was added are shown in lane 4 (T+G5/27) or lane 5 (C+G3/27). M, molecular weight markers in bp.

specific and saturable reaction, inhibited by a anti-EPR-1 mAb (Altieri, 1994). Therefore, a functional EPR-1 mRNA was apparently generated in CHO transfectants through the processing of a transcript containing the intervening sequence. Consistent with this possibility, Northern hybridization studies revealed that the EPR-1 message in CHO transfectants appeared as a prominent ~1.5 kb band, as opposed to the larger 1.9 kb transcript detected in mRNA or total RNA of EPR-1$^+$ HEL or PHA-activated T cells, under the same hybridizing conditions (Figure 3, left). In this context, the flanking oligonucleotide pair T+G5/27 that generated a 579 bp band in EPR-1$^+$ cells (Figure 2), amplified a considerably smaller DNA product of 128 bp from reverse-transcribed RNA of CHO transfectants (Figure 3, right). Similar results were obtained from reverse-transcribed RNA of chronic lymphocytic leukemia (CLL) lymphocytes, or from WEHI.3 monocytes, the latter using the more distantly spaced flanking oligonucleotide pair C+G3/27 (Figure 3, right).

DNA sequencing of the cloned PCR product amplified from WEHI.3 monocytes (PCR124) demonstrated a complete removal of the intervening sequence under these experimental conditions, thus generating a functionally uninterrupted transcript which joined the 5′ ORF in the EPR-1 mRNA with the 785 bp ORF downstream the intervening sequence (Figure 1A). A 5′ junction between the ORF and the intervening sequence was found at position 226 with the sequence CCCAGCCA:gtcccca, while a 3′ junction was found at position 677 with the sequence ctcagattcaacag:GCACCT (translated sequences are in capital letters). The identity of the EPR-1 intervening sequence as a retained intron and the authenticity of the 3′,5′ splice sites were independently substantiated in two separate series of experiments. First, two different oligonucleotides in the EPR-1 intervening sequence generated PCR products of expected sizes of 380 bp and 220 bp from a human genomic library (Figure 1, Supplementary Material). Second, a reverse "junctional" oligonucleotide GA3 spanning across the 3′ EPR-1 intron:exon boundary specifically amplified in combination with forward oligonucleotide C3/27 a single DNA product of 165 bp from reverse-transcribed RNA of HEL cells or CHO transfectants, consistent with the size of the functionally spliced EPR-1 transcript (Figure 2, Supplementary Material). The authenticity of the DNA product amplified by the "junctional" oligonucleotide GA3 and its identity with the functionally spliced EPR-1 transcript was independently validated both by Southern hybridization with an internal $^{32}$P-labeled EPR-1 oligonucleotide and by direct DNA sequencing of the cloned PCR product (Figure 2, Supplementary Material).

Cell type-specific removal of the intervening sequence directly correlated with surface expression of EPR-1. As shown in Figure 4, the reactivity of anti-EPR-1 mAb 2E1 with CLL lymphocytes or WEHI.3 monocytes was quantitatively 5−20-fold higher than that observed with other EPR-1-expressing cells, including Daudi and HEL (Figure 4), predominantly characterized by the unspliced EPR-1 transcript (Figure 2).

The alternative processing of EPR-1 mRNA was further quantitatively characterized. First, consistent with the data presented in Figure 2, amplification of reverse-transcribed mRNA of PHA-activated lymphocytes with the flanking oligonucleotide pair C+G3/27 generated a single DNA product of 702 bp (Figure 5, A) that was identical to the EPR-1 consensus sequence and contained the 451 bp intervening sequence (Figure 1A). Although no additional lower molecular weight bands with the expected size of 253 bp were visible by ethidium bromide staining (Figure 5A), Southern blot of the PCR product with a intronless EPR-1 probe (Figure 3) revealed two strong hybridization bands of 700 and 250 bp, respectively, the latter consistent with the size of the functionally spliced EPR-1 transcript (Figure 5, B). When interpreted in the context of competitive PCR studies (Gilliland et al., 1990), these data suggest a quantitative ratio ≤1:17 for the spliced vs unspliced EPR-1 mRNA.

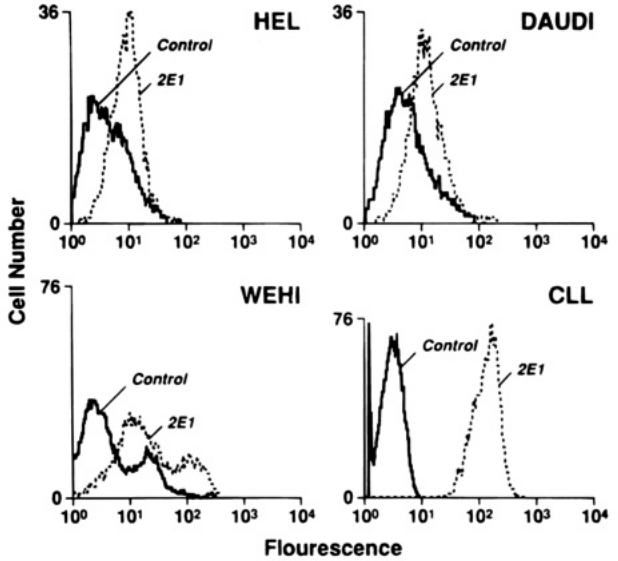


FIGURE 4: Removal of the intervening sequence correlates with EPR-1 surface expression. EPR-1-expressing cells were harvested, blocked with 20% normal human serum, and analyzed for their reactivity with anti-EPR-1 mAb 2E1 by flow cytometry. Background fluorescence was assessed in the presence of the irrelevant mAb HB3.

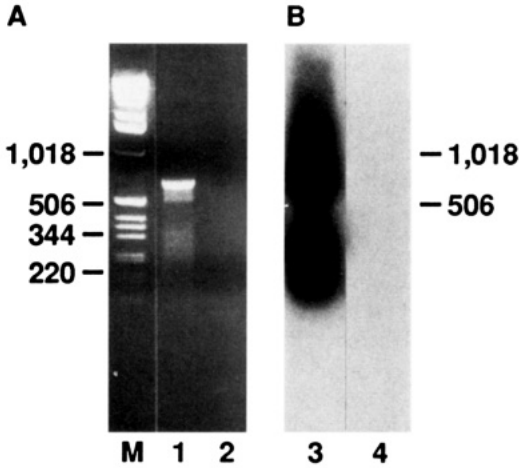


FIGURE 5: Amplification of two forms of EPR-1 mRNA from activated lymphocytes. (A) Ethidium bromide. A single DNA product of 702 bp containing the EPR-1 intervening sequence was amplified from reverse transcribed poly(A)$^+$ mRNA of PHA-activated lymphocytes by the flanking oligonucleotide pair C+G3/27 (lane 1). No additional lower molecular weight bands with the expected size of 253 bp were detected by ethidium bromide staining. (Lane 2) No template control incubation reaction. M, molecular weight marker in bp. (B) Southern hybridization. The PCR products shown in panel A were transferred to GeneScreen and hybridized with a $^{32}$P-labeled intronless EPR-1 probe (PCR106, 128 bp). Two radioactive bands of 700 and 250 bp (lane 3), consistent with the sizes of the unspliced (EPR-1b), and spliced (EPR-1a) mRNAs, respectively, were detected in the PCR products shown in lane 1. (Lane 4) Hybridization of control PCR products in lane 2.

Second, in ribonuclease protection studies of HEL or THP-1 cell RNA, the intronless EPR-1 probe PCR124 (see above) strongly hybridized with the predicted protected fragments of 146 and 107 bp of the unspliced EPR-1 transcript (Figure 6 and data not shown). In contrast, the functionally spliced EPR-1 mRNA was detected as a considerably fainter hybridization band of 253 bp, under the same experimental conditions (Figure 6 and data not shown). Densitometric quantitation of the autoradiograms from two independent experiments indicated that ~70% or ~93% of the EPR-1

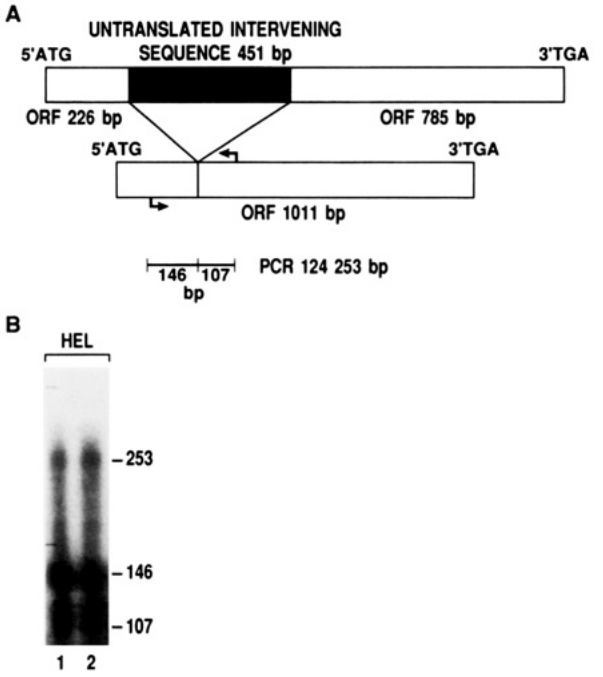


FIGURE 6: Ribonuclease protection assay. (A) Organization of unspliced/spliced EPR-1 mRNAs and relative position of the intronless probe PCR124. (B) HEL RNA (lane 1, 8 $\mu$g; lane 2, 24 $\mu$g) was hybridized with 4−8 × $10^4$ cpm of EPR-1 probe PCR124 (antisense strand) for 12 h at 45 °C. Protected fragments were separated by electrophoresis on a 5% acrylamide/8 M urea denaturing gel followed by autoradiography. Position of protected fragments of 253 bp, corresponding to the functionally spliced EPR-1a mRNA (A), or of 146 bp and 107 bp, corresponding to the unspliced EPR-1b transcript, are shown on the right.

mRNA retained the intervening sequence in HEL or THP-1 cells, respectively.

*Intron-Dependent Expression of EPR-1 in Stable Transfectants.* To investigate a potential requirement for intron-regulatory motif(s) in the processing of EPR-1 mRNA, CHO cells were transfected with the EPR-1 cDNA containing the intervening sequence ($\lambda$407), or with a synthetic minigene of the functionally spliced EPR-1 mRNA (F1). As shown in Figure 7, $\lambda$407 transfectants consistently displayed a 2−5-fold increased levels of EPR-1 surface expression, as compared with CHO cells transfected with the intronless F1 minigene (Figure 7). In control experiments, transfection of CHO cells with the intronless cDNA for ICAM-1 (Simmons et al., 1988) resulted in high levels of ICAM-1 surface expression (500 units of fluorescence in the 4-log horizontal axis shown in Figure 7), under the same experimental conditions (not shown).

*The Unspliced mRNA Generates a Novel Truncated Protein.* The translated product of the most prevalent, unspliced EPR-1 transcript is a truncated protein of 110 amino acids with a calculated $M_r$ of 14,150, structurally characterized by a unique, 34 residue-long -COOH terminus encoded by the first 104 bp of the intervening sequence (Table 1). Inspection of this polypeptide sequence, shown in Table 1, revealed a 8 amino acid motif, PPQHRAKS, which is highly homologous to the sequence PPQKKIKS found in N-*myc* (Stanton et al., 1986) and that functions as a specific nuclear targeting signal (Dang & Lee, 1989). To determine the genuine synthesis and potential subcellular localization of this variant form of EPR-1, designated EPR-1b, a sequence-specific antiserum was raised against the

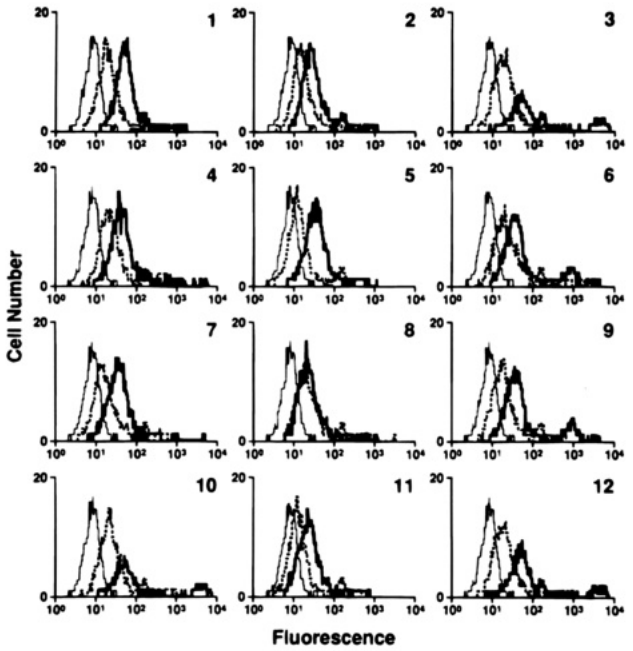


FIGURE 7: Intron-dependent expression of EPR-1 in stable transfectants. CHO cells were transfected with a EPR-1 cDNA construct containing the intervening sequence ($\lambda$407) or with a synthetic minigene of the functionally spliced EPR-1 mRNA (F1), selected, and cloned by limiting dilution. Flow cytofluorometric reactivity of anti-EPR-1 mAb 2E1 (Altieri, 1994) with WT CHO cells (solid lines), F1 transfectants (broken lines), or $\lambda$407 transfectants (bold lines), is shown for 12 representative independent clones.

Table 1: Translated Products Generated by Alternative Splicing of EPR-1 mRNA[a]

| | |
|---|---|
| MTSRGFQRSCNNPPCSSMTGRRANQIHHLTPDFSLRELLPPKKAGTW | EPR-1a, spliced |
| MTSRGFQRSCNNPPCSSMTGRRANQIHHLTPDFSLRELLPPKKAGTW | EPR-1b, unspliced |
| ADCVSPPCGERDRCEGWADRHTACSSPASTCQVHTQDCDSLNNMRS | EPR-1a, spliced |
| ADCVSPPCGERDRCEGWADRHTACSSPASPPQHRAKSFKKDLGHCL | EPR-1b, unspliced |
| 110 | |
| RHIHCGRLCHANKAVSSS→226 amino acids | EPR-1a, spliced |
| **FLREERPACAPRLTGPV Stop** | EPR-1b, unspliced |

[a] The consensus DNA sequence and corresponding protein translation was deduced from 32 independent clones isolated from four oligo-dT-primed human cDNA libraries. Cysteine residues are shadowed. The intron-encoded -COOH terminus of EPR-1b is in bold. A potential nuclear targeting signal in EPR-1b is boxed.

synthetic peptide P-25-A, duplicating the first 25 residues of the intron-encoded -COOH terminus of EPR-1b (Table 1). This antibody specifically immunoblotted a single band of ~14 kDa from detergent-solubilized THP-1 cell extracts (not shown). In indirect immunofluorescence staining of permeabilized HUVEC, the anti-P-25-A antiserum revealed a prominent nuclear reactivity, as compared with the background staining of control preimmune rabbit serum (Figure 8). Specificity for the observed nuclear reactivity of the anti-P-25-A antiserum was substantiated in control experiments using a anti-Sm nuclear protein antiserum, under the same experimental conditions (Figure 8).

## DISCUSSION

This study has shown that cell surface expression of a novel membrane receptor for factor Xa designated effector

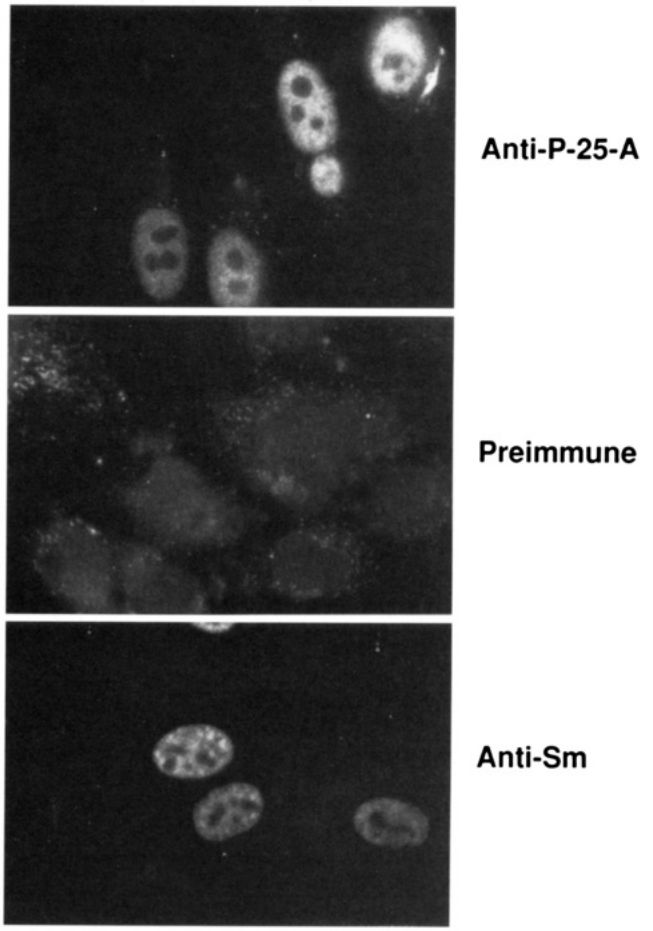


FIGURE 8: Nuclear localization of EPR-1b. A rabbit sequence-specific antiserum was raised against the synthetic peptide P-25-A, duplicating the first 25 residues of the intron-encoded -COOH terminus of the unspliced product of EPR-1 mRNA (EPR-1b). The anti-P-25-A antiserum showed a prominent nuclear reactivity in indirect immunofluorescence staining of permeabilized HUVEC (upper panel), as compared with background fluorescence observed with preimmune rabbit serum (middle panel). A anti-Sm nuclear protein antiserum was used as a control for nuclear localization staining, under the same experimental conditions (lower panel).

cell protease receptor-1 (EPR-1) (Altieri & Edgington, 1990; Worfolk et al., 1992; Altieri, 1994) is dynamically modulated by mRNA splicing. Differential processing of the EPR-1 transcript includes (i) cell type- and template-specific removal of an intervening sequence almost invariably retained in mature mRNA, and (ii) generation of a novel, nuclear-localized, truncated protein (EPR-1b) encoded by the most prevalent, unspliced mRNA.

Since the same organization of EPR-1 mRNA was found in 32 independent cDNA clones and was confirmed by amplification of reverse-transcribed mRNA of EPR-1$^+$ cells, it seems unlikely that this intervening sequence might represent a cloning artifact or reflect remnants of EPR-1 pre-mRNAs. Rather, several lines of evidence identify this region as an intron retained in the mRNA. First, its 5′,3′ junctional nucleotides GT/CAG meet the "invariant" consensus for intron boundaries (Shapiro & Senapathy, 1987; Jackson, 1991). Second, transfection of mammalian cells with the EPR-1 cDNA containing the intervening sequence was associated with mRNA processing, removal of the intervening sequence, and generation of a transcript directing cell surface expression of a functional factor Xa receptor (Altieri, 1994). Lastly, the EPR-1 intervening sequence was specifically amplified from human genomic DNA.

The intrinsically weak consensus at intron−exon boundaries may determine the ubiquitous retention of this intervening sequence in the mRNA. Statistical analysis of the EPR-1 junctional sequences (Shapiro & Senapathy, 1987) revealed a dramatic divergence from the consensus at the 5′ splice site, with a very low score of only 57%. Inspection of the 3′ splice site showed a higher score of 80%, although the rare presence (<2% of RNAs) of an AG dinucleotide within 10 bp from the splice site prevents additional comparisons (Shapiro & Senapathy, 1987). Such a divergence from the consensus is expected to drastically affect the process of splice site selection during spliceosome assembly (Green, 1991; Maniatis, 1991; Sharp, 1987), thus resulting in a very low splicing efficiency. Consistent with this scheme, conventional alignment algorithms do not identify the EPR-1 junctional sequences as potential intron:exon boundaries, leading to the speculation that a similar mechanism of gene regulation might unsuspectedly occur in other mammalian mRNAs.

While a cell type-specific pattern of mRNA processing (Mattox et al., 1993) is suggested by the identification of certain cells that constitutively overexpress EPR-1, i.e., CLL lymphocytes, the intervening sequence itself may contain regulatory signal(s) that contribute to the splicing process. This is underscored by the ability of a cDNA construct containing the intervening sequence to consistently direct higher levels of EPR-1 surface expression in stably transfected cells, as compared with transfectants carrying a intronless EPR-1 minigene. A role for discrete intron *cis*-acting sequences in regulating gene expression has been previously reported for cell-cycle dependent mechanisms of posttranscriptional control (Ash et al., 1993), alternative splicing of viral gene products (Brady et al., 1992), and in vitro polyadenylation (Niwa et al., 1990).

Regulation of gene expression by intron retention modulates adenovirus and cytomegalovirus gene products (Dix & Leppard, 1993; Stenberg et al., 1985), or sex-determining gene(s) and P-transposable elements of *Drosophila* (Laski et al., 1986; Bell et al., 1988). However, there are two important features that distinguish the model of intron splicing described here from previous examples of intron-dependent gene regulation in mammalian cells. First, at variance with bovine growth hormone (Hampson et al., 1989) or decay accelerating factor (Caras et al., 1987), in which only 2−10% of the mRNA retains an intron, the unspliced EPR-1 transcript constitutes approximately 70−90% of the mature mRNA. Second, differential processing of EPR-1 mRNA introduces a profound protein isoform diversity. Instead of EPR-1a, a transmembrane glycoprotein receptor for factor Xa (Altieri, 1994), the product of the very abundant and widespread unspliced EPR-1 transcript is a novel truncated protein of 110 amino acids, designated EPR-1b, characterized by a unique intron-encoded -COOH terminus containing the sequence PPQHRAKS. The highly homologous motif PPQKKIKS found in the -COOH terminus of N-*myc* (Stanton et al., 1986) is per se sufficient to direct nuclear localization of a reporter gene, thus functioning as a genuine nuclear targeting signal (Dang & Lee, 1989). In this context, a intron sequence-specific antiserum confirmed the existence of EPR-1b and revealed its prominent nuclear localization in permeabilized HUVEC. Although the potential function(s) of this variant isoform are currently unknown, its structural and topographical features suggest a role unrelated to cellular responses initiated by occupancy of EPR-1a with factor Xa on the cell surface (Altieri & Edgington, 1990; Worfolk et al., 1992; Altieri & Stamnes, 1994).

Although cellular protease receptors for thrombin (Vu et al., 1991), urokinase (Roldan et al., 1990), or factor Xa (Altieri, 1994) participate in pleiotropic signaling pathways, little is known about how protease-dependent cellular responses are regulated. In this context, modulation of EPR-1a surface expression by mRNA splicing might provide a mechanism to regulate receptor-mediated cellular responses initiated by factor Xa. These include membrane assembly of multiple proteolytic activities (Worfolk et al., 1992), proliferation of vascular smooth muscle cells (Gasic et al., 1992), and lymphocyte subsets (Altieri and Stamnes, 1994), and the ability to release endothelial cell mitogens (Gajdusek et al., 1986). That EPR-1a expression can be dynamically modulated has been anticipated previously, with the demonstration of an 8−10-fold increased membrane density of EPR-1a during lymphocyte activation and proliferation in vitro (Altieri & Edgington, 1990). In summary, these studies underscore a broader participation of intron retention mechanisms in modulating mammalian gene expression and suggest a novel paradigm for the regulation of factor Xa-dependent cellular effector functions.

## ACKNOWLEDGMENT

The author thanks Drs. B. Seed, R. O. Hynes, P. A. Sharp, and S. Berget for helpful discussion, A. Duperray for transfections, and A. Almenar for indirect immunofluorescence. The expert technical assistance of Ms. S. Stamnes and A. Lewis is kindly acknowledged.

## SUPPLEMENTARY MATERIAL AVAILABLE

Two figures showing genomic amplification of the EPR-1 intervening sequence and amplification of EPR-1 sequences with a splice-site-specific oligonucleotide (3 pages). Ordering information is given on any current masthead page.

## REFERENCES

Altieri, D. C. (1994) *J. Biol. Chem. 269*, 3139−3142.
Altieri, D. C., & Edgington, T. S. (1990) *J. Immunol. 145*, 246−253.
Altieri, D. C., & Stamnes, S. J. (1994) *Cell. Immunol. 155*, 372−383.
Ash, J., Ke, Y., Korb, M., & Johnson, L. F. (1993) *Mol. Cell. Biol. 13*, 1565−1571.
Bar-Shavit, R., Kahn, A., Wilner, G. D., & Fenton, J. W., II. (1983) *Science 220*, 728−731.
Bell, L. R., Maine, E. M., & Cline, T. W. (1988) *Cell 55*, 1037−1046.
Brady, H. A., Scaria, A., & Wold, W. S. (1992) *J. Virol. 66*, 5914−5923.
Caras, I. W., Davitz, M. A., Rhee, L., Weddel, G., Martin, D. W., Jr., & Nussenzweig, V. (1987) *Nature 325*, 545−549.
Dang, C. V., & Lee, W. M. F. (1989) *J. Biol. Chem. 264*, 18019−18023.
Daniel, T. O., Gibbs, V. C., Milfay, D. F., Garavoy, M. R., & Williams, L. T. (1986) *J. Biol. Chem. 261*, 9579−9582.
Dix, I., & Leppard, K. N. (1993) *J. Virol. 67*, 3226−3231.
Furie, B., & Furie, B. C. (1988) *Cell 53*, 505−518.

Gajdusek, C., Carbon, S., Ross, R., Nawroth, P., & Stern, D. (1986) *J. Cell Biol. 103*, 419−428
Gasic, G. P., Arenas, C. P., Gasic, T. B., & Gasic, G. J. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 2317−2320.
Gilliland, G., Perrin, S., Blanchard, K., & Bunn, H. F. (1990) *Proc.Natl.Acad.Sci. U.S.A. 87*, 2725−2729.
Glenn, K. C., & Cunningham, D. D. (1979) *Nature 278*, 711−714.
Green, M. R. (1991) *Annu. Rev. Cell Biol. 7*, 559−599.
Hampson, R. K., La Follette, L., & Rottman, F. M. (1989) *Mol. Cell. Biol. 9*, 1604−1610.
Jackson, I. J. (1991) *Nucleic Acids Res. 19*, 3795−3798.
Kirchheimer, J. C., Wojita, J., Christ, G., & Binder, B. R. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5424−5428.
Kozak, M. A. (1984) *Nucleic Acids Res. 12*, 857−872.
Laski, F. A., Rio, D. C., & Rubin, G. M. (1986) *Cell 44*, 7−19.
Maniatis, T. (1991) *Science 251*, 33−34.
Mattox, W., Ryner, L., & Baker, B. S. (1992) *J. Biol. Chem. 267*, 19023−19026.
Niwa, M., Rose, S. D., & Berget, S. M. (1990) *Genes Dev. 4*, 1552−1559.
Ossowski, L. (1987) *Cell 107*, 2437−2445.
Roldan, A. L., Cubellil, M. V., Masucci, M., Behrendt, N., Lund, L. R., Dano, K., Appella, E., & Blasi, F. (1990) *EMBO J. 9*, 467−474.
Shapiro, M. B., & Senapathy, P. (1987) *Nucleic Acid Res. 15*, 7155−7174.
Sharp, P. A. (1987) *Science 235*, 766−771.
Simmons, D., Makgoba, M. W., & Seed, B. (1988) *Nature 331*, 624−627.
Smith, C. W. J., Patton, J. G., & Nadal-Ginard, B. (1989) *Annu. Rev. Genet. 23*, 527−577.
Stanton, L. W., Schwab, M., & Bishop, J. M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1772−1776.
Stenberg, R. M., Witte, P. R., & Stinski, M. F. (1985) *J. Virol. 56*, 665−675.
Vu, T. K. H., Hong, D. T., Wheaton, V. I., & Coughlin, S. R. (1991) *Cell 64*, 1057−1068.
Worfolk, L. A., Robinson, R. A., & Tracy, P. B. (1992) *Blood 80*, 1989−1997.