MOLECULAR AND CELLULAR BIOLOGY, Dec. 1989, p. 5718–5721
0270-7306/89/125718-04$02.00/0

Vol. 9, No. 12

Definition of an Ets1 Protein Domain Required for Nuclear Localization in Cells and DNA-Binding Activity In Vitro

K. E. BOULUKOS,† P. POGNONEC,‡ B. RABAULT, A. BEGUE, AND J. GHYSDAEL*

Institut National de la Santé et de la Recherche Medicale U186/Centre National de la Recherche Scientifique UA 041160, Institut Pasteur, 1 rue Calmette, 59019 Lille, France

Received 17 July 1989/Accepted 3 September 1989

Ets1 and Ets2 are nuclear phosphoproteins which bind to DNA in vitro and share two domains of strong identity. Deletion analyses of each of these conserved regions in Ets1 demonstrated that integrity of the carboxy-terminal domain, also conserved in the more distantly related *elk* and *erg* gene products, is essential for both nuclear targeting and DNA-binding activity in vitro.

The c-*ets* family consists of genes with sequence identity to the v-*ets* oncogene found in the genome of the transforming E26 retrovirus. The c-*ets* family includes c-*ets-1* (5, 12, 21), the progenitor of v-*ets* (6, 8, 10, 13), c-*ets-2* (3, 11, 21, 22), and the more distantly related *erg* and *elk* genes (16, 17, 19). Although the functions of these genes remain to be established, recent evidence suggests a role for the c-*ets-1* product (Ets1) in lymphoid activation (14), whereas the c-*ets-2* product (Ets2) may be involved in cell proliferation (2) and in macrophage functional activation (K. E. Boulukos, P. Pognonec, E. Sariban, M. Bailly, C. Lagrou, and J. Ghysdael, submitted for publication).

Comparison of the amino acid sequence of chicken Ets1 and Ets2 allowed the identification of two regions of strong identity of 96 and 175 amino acids localized near the amino terminus (domain B in Fig. 1) and in the carboxy-terminal domains (domains D and E in Fig. 1), respectively. This last domain can be divided in two regions of 50% identity (region D) and more than 90% identity (region E). The D domain is separated from the B domain by a 113-amino-acid domain (domain C) displaying no identity with the corresponding region of Ets2. Finally, the extreme amino-terminal domain of Ets1 (domain A) shares virtually no identity with Ets2 and, in fact, corresponds to a region where alternative splicing events are known to occur in both chicken c-*ets-1* and c-*ets-2* genes (13; Boulukos et al., submitted). Since both Ets1 and Ets2 are nuclear phosphoproteins (3, 9, 14, 15) and bind to DNA in vitro (15; Boulukos et al., submitted), we hypothesized that the regions of strong homology were the most likely to be responsible for the nuclear localization and DNA-binding activity in vitro of these proteins. To test this, we constructed a series of mutants (Fig. 1) using a cDNA clone encoding Ets1 (15). The resulting mutated DNAs were subcloned into a simian virus 40 (SV40) early promoter-based expression vector (4) and transfected into monkey COS-1 cells. Two days after transfection, the subcellular locations of mutant Ets1 proteins were analyzed by indirect immunofluorescence with an Ets-specific antiserum and compared with the location of a wild-type protein similarly expressed in COS-1 cells from the previously described pKCR3-c-*ets-1* construct (15). Similarly, the DNA-binding abilities of mutant Ets1 proteins were compared with that of wild-type Ets1 by DNA-cellulose chromatography of extracts of L-[^{35}S]methionine-labeled cells and immunoprecipitation analyses with an Ets-specific antiserum as described previously (15).

The first of the mutants constructed involved a 279-base-pair internal in-frame deletion from *Hpa*I at nucleotide (nt) position 110 to *Bgl*II at nt position 389 (for map coordinates, see reference 8), resulting in the deletion of almost all of domain B of Ets1 (Δ38-130). A high level of immunofluorescence was detected in the nuclei of COS-1 cells transfected with the Δ38-130 construct (Fig. 2, Δ38-130). The pattern of immunofluorescence derived from the Δ38-130 mutant was indistinguishable from that observed in COS-1 cells transfected with a wild-type Ets1 construct (Fig. 2, WT). This nuclear immunofluorescence is specific to Ets1 since it was not detected with either a nonimmune control serum or an Ets antiserum blocked with an excess of the purified bacterial immunogen (data not shown). These results demonstrate that domain B of Ets1 can be deleted without affecting the subcellular localization of the protein.

We next constructed a series of deletion mutants in the 3′ region of the c-*ets-1*-specific cDNA clone corresponding to the highly conserved carboxy-terminal domain E. To delete the first 138 nt of domain E, we created a novel *Hin*dIII site by point mutating a C to a T at nt position 955 using an oligonucleotide-directed in vitro mutagenesis system (Amersham Corp., Arlington Heights, Ill.). This point mutation resulted in an amino acid change from a proline to a leucine at amino acid 319 (319P-L). Immunofluorescence experiments demonstrated that the product of the 319P-L construct remains localized in the nucleus, indicating that this amino acid change does not affect the subcellular localization of the protein (Fig. 2, 319P-L). The creation of this novel *Hin*dIII site allowed us to remove 46 amino acids of conserved sequences of domain E by digestion of the 319P-L construct with restriction enzyme *Hin*dIII followed by religation, thereby generating the in-frame deletion mutant Δ318-363. The Δ318-363 protein entirely localized in the nucleus (Fig. 2, Δ318-363). These results demonstrated that integrity of the first 46 amino acids of domain E is not required for nuclear localization of Ets1 in cells. We therefore constructed another mutant, Δ364-441, deleting all nt downstream of the natural *Hin*dIII site found at nt position 1092 (numbering as reported in reference 8) in the c-*ets-1* cDNA. Immunofluorescence analysis of COS-1 cells transfected with this mutant showed that Δ364-441 Ets1 is exclusively localized in the

* Corresponding author.
† Present address: Department of Biochemistry, New York University Medical Center, New York, NY 10016.
‡ Present address: Department of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY 10021.

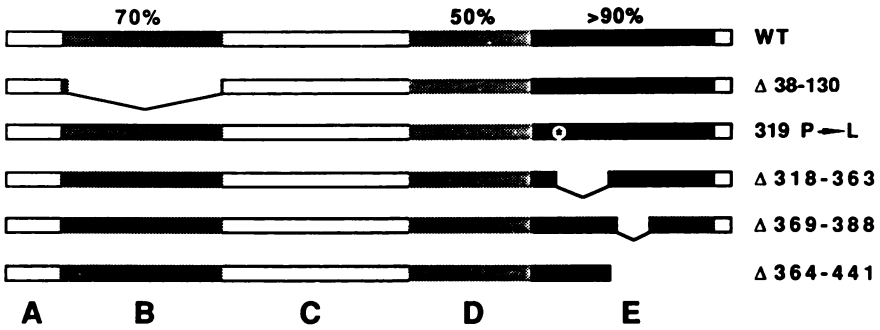

FIG. 1. Wild-type and mutant c-*ets-1* constructions. The first line drawing illustrates the wild-type Ets1. All other line drawings schematically depict the mutant constructions used in this study. Light grey boxes indicate amino acid sequence identities of approximately 70 and 50% shared between Ets1 and Ets2. The dark shaded box represents amino acid sequence identity greater than 90% shared between Ets1 and Ets2. Δ38-130 was constructed with a c-*ets-1* cDNA clone that was digested with *Bgl*II restriction enzyme and subsequently blunt ended with Klenow enzyme. After digestion with restriction enzyme *Hpa*I, the blunt-ended *Bgl*II site was ligated in phase to the *Hpa*I site with T4 ligase. The circularized mutated molecule was then digested with *Eco*RI and subcloned into the *Eco*RI site of an SV40 promoter-based expression vector (4). Similarly, Δ364-441 was constructed by digesting the c-*ets-1* cDNA clone with *Hin*DIII and *Eco*RI, thereby removing the C-terminal-coding sequences downstream of the true *Hin*DIII site. This mutated DNA was then cloned into a similar SV40-based expression vector containing a *Hin*DIII site upstream from a stop codon. The point mutation at nt position 955 creating a novel *Hin*DIII site in a c-*ets-1* cDNA clone (319P-L) as well as Δ369-388 were generated by oligonucleotide-directed mutagenesis (Amersham). Δ318-363 was obtained by *Hin*DIII digestion and religation of mutant 319P-L DNA. Abbreviations: WT, wild type; Δ38-130, deletion in domain B where, as in all constructions, Δ represents deletion and the first number which follows Δ corresponds to the first amino acid deleted and the second number corresponds to the last amino acid deleted; 319P-L, the amino acid at position 319 (asterisk) has been changed from a proline to a leucine; Δ318-363, deletion in the first 46 amino acids of domain E; Δ369-388, deletion including the eight-basic-amino-acid motif; Δ364-441, deletion of the 78 C-terminal amino acid residues of domain E.

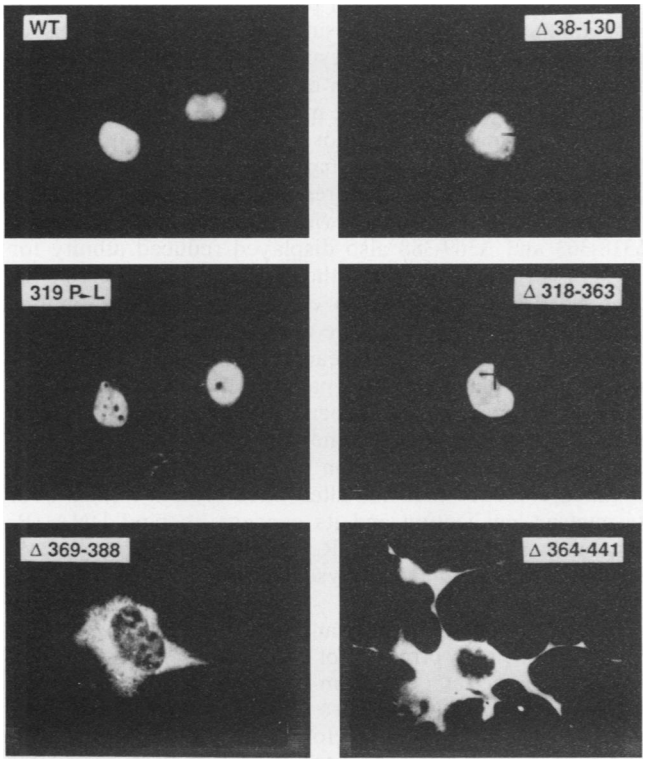

FIG. 2. Immunofluorescence studies. Monkey COS-1 cells were transfected with 10 μg of SV40-based expression vector encoding either wild-type (WT) or mutant DNA as indicated essentially as previously described (3, 15). Briefly, cells were transferred to collagen-coated cover slips and fixed in cold methanol (−20°C) for 20 min. After extensive washes in phosphate-buffered saline and incubation in ethanolamine, subcellular Ets1 localization was analyzed by immunofluorescence with our standard Ets antiserum (12). After being washed, cells were incubated with fluorescein isothiocyanate-conjugated swine anti-rabbit immunoglobulins (Dakopatts, Glostrup, Denmark). Immunofluorescence was examined with a Zeiss epifluorescence microscope. The experiments detailed here gave a consistent and reproducible immunofluorescence pattern in >95% of exogenous Ets1-expressing cells.

cytoplasm (Fig. 2, Δ364-441). This indicated that integrity of the 78 C-terminal residues of Ets1 is required for nuclear localization of the protein. As we previously noted (15), a stretch of basic amino acid residues (376-GKRKNKPK) is found within this carboxy-terminal domain and is similar to the minimal amino acid sequence important for nuclear targeting of SV40 large T antigen as well as several other nuclear proteins (7, 20). In addition, this region is predicted to extend out of the surface of the protein (15). We therefore deleted these basic sequences (Δ369-388) using an oligonucleotide-directed in vitro mutagenesis system (Amersham). Mutant Δ369-388 Ets1 was almost exclusively localized in the cytoplasm of COS-1 transfected cells (Fig. 2, Δ369-388). These results indicated that these basic sequences represent an essential determinant for Ets1 nuclear localization. However, we reproducibly noticed a very low level of patchy nuclear immunofluorescence in COS-1 cells expressing the Δ369-388 protein. This residual nuclear immunofluorescence, which differs from the usual uniform pattern of immunofluorescence of wild-type Ets1, is not likely to be due to passive transfer of the smaller-sized mutated Δ369-388 protein since the Δ364-441 protein, which is even smaller in size, remained localized in the cytoplasm. This suggests that integrity of additional sequences located downstream of amino acid 388 is required to induce complete nuclear localization of Ets1. Similar observations have been made recently for several other nuclear proteins including Myc and polyomavirus large T (7, 20).

These Ets1 mutant constructions were next used to delineate the region(s) important for binding to sequence-nonspecific DNA in vitro. Wild-type and mutant constructions were

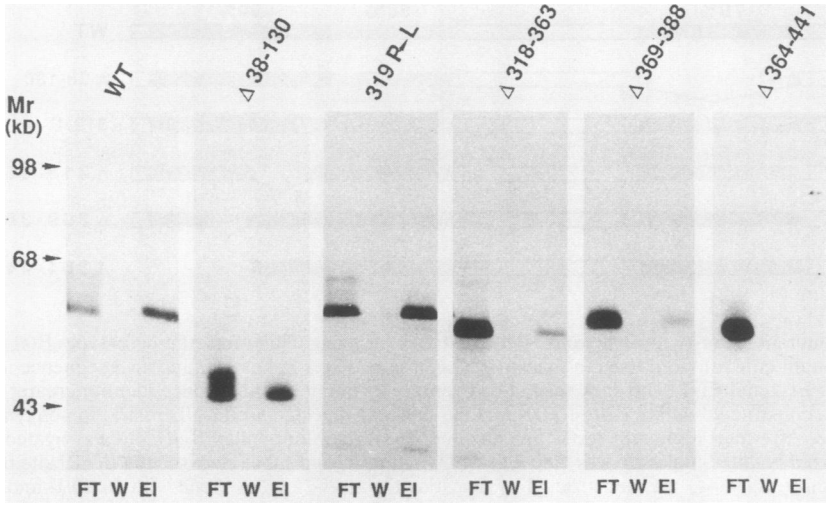

FIG. 3. DNA binding of wild-type Ets1 (WT) and mutant Ets1 in vitro. Monkey COS-1 cells were transfected with 10 μg of SV40-based expression vector encoding either wild-type or mutant Ets1 proteins. Two days after transfection, cells were labeled with 500 μCi of L-[^{35}S]methionine per ml and lysed in TDT buffer (10 mM Tris hydrochloride [pH 7.4], 50 mM NaCl, 1 mM dithiothreitol, 0.1% Triton X-100, 1 mM EDTA, 10 μg of leupeptin per ml, 100 μg of phenylmethylsulfonyl fluoride per ml, 1% aprotinin; all from Sigma Chemical Co., St. Louis, Mo.). After centrifugation at 100,000 × *g*, the supernatant was loaded on a native calf thymus DNA-cellulose column equilibrated in TDT buffer as described previously (15). After extensive washing in TDT buffer, bound proteins were eluted in TDT buffer adjusted to 1 M NaCl. Equivalent amounts of flowthrough (FT), wash (W), and elution (El) fractions were immunoprecipitated with an excess of the standard Ets antiserum and analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. kD, Kilodaltons.

transfected into COS-1 cells, the cells were labeled with L-[^{35}S]methionine and lysed, and lysates were incubated in the presence of calf thymus DNA-cellulose as previously described (15). The Δ38-130 protein maintained its ability to bind to DNA-cellulose as efficiently as the wild-type protein (WT) (Fig. 3, compare WT and Δ38-130 lanes). Therefore, although conserved 70%, domain B does not appear to be important for nuclear localization of Ets1 or for its ability to bind to DNA. As previously demonstrated, the slower-migrating high-molecular-weight species immunoprecipitated by the Ets antiserum are phosphorylated versions of Ets1 (14). Consistent with our recent observation with wild-type Ets1 (15), the phosphorylated forms of Δ38-130 were detected exclusively in the flowthrough fraction (Fig. 3, FT), whereas only the nonphosphorylated form bound to DNA-cellulose (Fig. 3, El). The 319P-L mutant protein with the amino acid change from a proline to a leucine residue also bound to DNA-cellulose as efficiently as wild-type Ets1 (Fig. 3, compare lanes WT and 319P-L, FT vs. El). The 3′ constructions with deletion of either the first 46 amino acids of domain E, Δ318-363, or the basic sequences, Δ369-388, were able to bind to DNA, but the resulting mutated proteins had a lower affinity for DNA than the wild-type protein did (Fig. 3). However, deletion of the 78 C-terminal residues of Ets1 in mutant Δ364-441 resulted in a complete loss of affinity for DNA-cellulose. Results similar to those depicted in Fig. 3 were obtained when extracts containing the wild-type or Δ364-441 protein were mixed before cochromatography on the same DNA-cellulose column (data not shown). These results demonstrated that within domain E, integrity of the extreme C-terminal region of Ets1 is absolutely required for DNA binding in vitro (Table 1). Since mutants Δ318-363 and Δ369-388 also displayed reduced affinity for DNA in vitro, at least part of the first 46 amino acid residues of domain E also appear to contribute to efficient DNA binding. Studies reported here did not address the possible role of domain D in Ets1 nuclear targeting and DNA-binding activity. This domain forms part of an entire exon in the chicken and human c-*ets-1* locus (10, 23). In mammals, an alternative splicing event generates a protein lacking the domain encoded by this exon (18), and we have recently obtained evidence that this alternative protein retains both its nuclear localization and its capacity to bind DNA (P. Pognonec, K. E. Boulukos, R. Bosselut, C. Boyer, A. M. Schmitt-Verhulst, and J. Ghysdael, submitted for publication).

TABLE 1. Summary of subcellular localization and DNA binding analyses of wild-type and mutant Ets1 constructions[a]

Construction	Localization	DNA binding
WT	N	+++
Δ38-130	N	+++
319 P-L	N	+++
Δ318-363	N	+
Δ369-388	C/N	+
Δ364-441	C	−

[a] Abbreviations: WT, wild type; N, nucleus; C, cytoplasm; +++, high affinity for DNA-cellulose; +, low affinity for DNA-cellulose; −, complete loss of affinity for DNA-cellulose.

The DNA-binding activity analyzed here is likely to represent an intrinsic property of Ets1 since an Ets1 protein produced in insect cells from a baculovirus recombinant vector also binds DNA in vitro (J. Ghysdael and D. Gijsen, manuscript in preparation). However, the predicted amino acid sequences of Ets1 and Ets2 in domain E do not reveal homologies to any of the known protein DNA-binding motifs. Recently, Anton and Frampton (1) proposed that tryptophan repeats represent a novel DNA-binding motif in Myb and Ets proteins. As they pointed out, three tryptophan repeats are found in domain E spaced at 17 and 18 amino acids. It is interesting that the first two of these tryptophans are deleted in Δ318-363 and the third tryptophan is deleted in Δ369-388 and that both these mutant Ets1 products exhibit a

reduced affinity for DNA. However, when the third tryptophan is deleted in a larger deletion (Δ364-441), the mutant Ets1 products completely lose the ability to bind to DNA-cellulose, suggesting that either the tryptophan spacing in itself is insufficient in functioning as a DNA-binding motif or possibly the tertiary structure of the protein is modified such that the remaining two tryptophans are rendered inaccessible, thereby disrupting the DNA-binding motif. Since domain E is highly conserved in all members of the *ets* family, including the more distantly related *erg* and *elk* gene products, our results suggest that this region is probably responsible for nuclear localization and DNA-binding activity of all members of the *ets* gene family, whereas other, less conserved regions of these protein products could be responsible for other aspects of *ets* protein function.

This work was supported by funds from the Institut National de la Santé et de la Recherche Medicale (INSERM), Centre National de la Recherche Scientifique (CNRS), Institut Pasteur de Lille, and Association pour la Recherche sur le Cancer (ARC).

LITERATURE CITED

1. **Anton, I. A., and J. Frampton.** 1988. Tryptophans in *myb* proteins. Nature (London) **336**:719.
2. **Bhat, N. K., R. J. Fisher, S. Fujiwara, R. Ascione, and T. S. Papas.** 1987. Temporal and tissue-specific expression of mouse *ets* genes. Proc. Natl. Acad. Sci. USA **84**:3161–3165.
3. **Boulukos, K. E., P. Pognonec, A. Begue, F. Galibert, J. C. Gesquière, D. Stehelin, and J. Ghysdael.** 1988. Identification in chickens of an evolutionarily conserved cellular *ets-2* gene (*c-ets-2*) encoding nuclear proteins related to the products of the *c-ets* proto-oncogene. EMBO J. **7**:697–705.
4. **Breathnach, R., and B. A. Harris.** 1983. Plasmids for the cloning and expression of full-length double-stranded cDNAs under the control of the SV40 early or late gene promoter. Nucleic Acids Res. **11**:7119–7136.
5. **Chen, J. H.** 1985. The proto-oncogene c-*ets* is preferentially expressed in lymphoid cells. Mol. Cell. Biol. **5**:2993–3000.
6. **Chen, J. H.** 1988. Complementary DNA clones of chicken proto-oncogene *c-ets*: sequence divergence from the viral oncogene *v-ets*. Oncogene Res. **2**:371–384.
7. **Dang, C. V., and W. M. F. Lee.** 1988. Identification of the human c-*myc* protein nuclear translocation signal. Mol. Cell. Biol. **8**:4048–4054.
8. **Duterque-Coquillaud, M., D. Leprince, A. Flourens, C. Henry, J. Ghysdael, B. Debuire, and D. Stehelin.** 1988. Cloning and expression of chicken $p54^{c\text{-}ets}$ cDNAs: the first $p54^{c\text{-}ets}$ coding exon is located into the 40 kbp genomic domain unrelated to *v-ets*. Oncogene Res. **2**:335–344.
9. **Fujiwara, S., R. J. Fisher, A. Seth, N. K. Bhat, S. D. Showalter, M. Zweig, and T. S. Papas.** 1988. Characterization and localization of the products of the human homologs of the *v-ets* oncogene. Oncogene **2**:99–103.
10. **Gegonne, A., D. Leprince, M. Duterque-Coquillaud, B. Vandenbunder, A. Flourens, J. Ghysdael, B. Debuire, and D. Stehelin.** 1987. Multiple domains for the chicken cellular sequences homologous to the *v-ets* oncogene of the E26 retrovirus. Mol. Cell. Biol. **7**:806–812.
11. **Ghysdael, J., A. Gegonne, P. Pognonec, K. Boulukos, D. Leprince, D. Dernis, C. Lagrou, and D. Stehelin.** 1986. Identification in chicken macrophages of a set of proteins related to, but distinct from, the chicken cellular *c-ets*-encoded protein $p54^{c\text{-}ets}$. EMBO J. **5**:2251–2256.
12. **Ghysdael, J., A. Gegonne, P. Pognonec, D. Dernis, D. Leprince, and D. Stehelin.** 1986. Identification and preferential expression in thymic and bursal lymphocytes of a *c-ets* oncogene encoded Mr 54,000 cytoplasmic protein. Proc. Natl. Acad. Sci. USA **83**:1714–1718.
13. **Leprince, D., M. Duterque-Coquillaud, R. P. Li, C. Henry, A. Flourens, B. Debuire, and D. Stehelin.** 1988. Alternative splicing within the chicken c-*ets-1* locus: implications for transduction within the E26 retrovirus of the c-*ets* proto-oncogene. J. Virol. **62**:3233–3241.
14. **Pognonec, P., K. E. Boulukos, J. C. Gesquière, D. Stéhelin, and J. Ghysdael.** 1988. Mitogenic stimulation of thymocytes results in the calcium-dependent phosphorylation of *c-ets-1* proteins. EMBO J. **7**:977–983.
15. **Pognonec, P., K. E. Boulukos, and J. Ghysdael.** 1989. The *c-ets-1* protein is chromatin associated and binds to DNA in vitro. Oncogene **4**:691–697.
16. **Rao, V. N., K. Huebner, M. Isobe, A. Ar-Rushdi, C. M. Croce, and E. S. P. Reddy.** 1989. *elk*, tissue-specific *ets*-related genes on chromosomes X and 14 near translocation breakpoints. Science **244**:66–70.
17. **Rao, V. N., T. S. Papas, and E. S. P. Reddy.** 1987. *erg*, a human *ets*-related gene on chromosome 21: alternative splicing, polyadenylation, and translation. Science **237**:635–639.
18. **Reddy, E. S., and V. N. Rao.** 1989. Structure, expression and alternative splicing of the human *c-ets1* protooncogene. Oncogene Res. **3**:239–246.
19. **Reddy, E. S. P., V. N. Rao, and T. S. Papas.** 1987. The *erg* gene: a human gene related to the *ets* oncogene. Proc. Natl. Acad. Sci. USA **84**:6131–6135.
20. **Richardson, W. D., B. L. Roberts, and A. E. Smith.** 1986. Nuclear localization signals in polyoma virus large-T. Cell **44**:77–85.
21. **Watson, D. K., M. J. McWilliams-Smith, M. F. Nunn, P. H. Duesberg, S. J. O'Brien, and T. S. Papas.** 1985. The *ets* sequence from the transforming gene of avian erythroblastosis. Proc. Natl. Acad. Sci. USA **82**:7294–7298.
22. **Watson, D. K., M. J. McWilliams, P. Lapis, J. A. Lautenberger, C. W. Schweinfest, and T. S. Papas.** 1988. Mammalian *ets-1* and *ets-2* genes encode highly conserved proteins. Proc. Natl. Acad. Sci. USA **85**:7862–7866.
23. **Watson, D. K., M. J. McWilliams, and T. S. Papas.** 1988. Molecular organization of the chicken *ets* locus. Virology **164**: 99–105.