Vol. 210, No. 1, 1995 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
May 5, 1995 Pages 181-188

# IDENTIFICATION OF A NOVEL REGULATORY ELEMENT IN THE C-MOS LOCUS THAT ACTIVATES TRANSCRIPTION IN SOMATIC CELLS

J.L. Lenormand and S.A. Leibovitch*

Laboratoire de Génétique Oncologique URA 1967 CNRS

Institut Gustave Roussy, 39, rue Camille Desmoulins 94800

Villejuif, FRANCE

Received March 28, 1995

We have identified a DNA sequence in the 5' flanking sequence of rat *c-mos* gene which fulfills operational criteria for enhancers, increasing transcription from heterologous promoters in somatic cells. This new enhancer region contains some specific motifs such as two CArG boxes, two M-CAT binding sites and CCAAT consensus sequences.This Upstream Enhancer region (UER) is recognized by distinct protein complexes and particularly CArG2 and M-CAT R1 motifs, which are adjacent in the DNA sequence. Several lines of evidence indicate that none of the two CArG boxes bind to the Serum response factor (SRF). Site-directed mutations of both the CArG2 and M-CAT R1 binding sites suppress their enhancer activity. These results suggest that direct and indirect interactions involving multiple nuclear factors and distinct elements of the UER may be required for its enhancer functions in somatic cells. 

**INTRODUCTION:** Earlier studies of *c-mos* expression argued for a specific role in maturation of germinal cells and led to the demonstration that in Xenopus eggs (1, 2) and in mouse oocyte system (3-5), $p39^{c\text{-}mos}$ mimics the cytostatic factor (CSF) to induce mitotic arrest (6). Various somatic tissues and cell lines were found to express detectable levels of *c-mos* transcripts and proteins (7-10) but little is known about the function(s) of *c-mos*. However the high expression of *c-mos* protein and various $c\text{-}mos^{xe}$ mutants with "stabilizating" penultimate residues in somatic cells induce the cell growth retardation or arrest and/or in some cases, death (11-15). Moreover, we and different laboratories have recently shown that $p34^{cdc2}$, a component of the maturation promoting factor (MPF) (16, 17) or $p35^{cdk}$ (18) were complexed with $c\text{-}mos^{rat}$, $c\text{-}mos^{xe}$ or *v-mos* proteins either in rat skeletal muscle or in transformed NIH3T3 cells. This original association between $p34^{cdc2}$ and $c\text{-}mos^{rat}$ protein in adult skeletal muscle pleads in favor of a role as cytostatic factor for the *c-mos* gene.

In rat myogenic cells and skeletal tissues two *c-mos* transcripts of 1.7kb and 3.6kb long are expressed from the same genomic locus (19). The minor 1.7kb species is similar to the *c-mos* RNA observed in rat ovary and in male germ cells (4, 10, 20). Recently, a

*: to whom correspondence should be addressed.

0006-291X/95 $5.00

3.6kb *c-mos* RNA has been detected in rat pachytene spermatocytes and early spermatids (20). Finally in a recent report, 1.7kb and 3.5kb *c-mos* messager RNAs were detected in various human cell lines (16). Expression of *c-mos* RNA in mouse oocytes, rat testes and rat skeletal muscle was controled by three distinct promoters in position 0.05kb (21), 0.56kb (20) and 1.010kb (22) with respect to the first ATG of the ORF (19). In rat skeletal muscle like in mouse oocyte, the tissue-specific promoter contains an Initiator element involved in the *c-mos* RNA expression (21, 22). The present study was undertaken in order to gain some insight on the transcriptional regulation of *c-mos* in somatic cells. We demonstrate that beside the sole enhancer-like element previously depicted (23, 24), an other regulatory upstream element (named UER for Upstream Enhancer Region) allows to activate heterologous promoters in different cell lines. This *c-mos* upstream element possesses two CArG-boxes, a consensus motif that is known to play a role either in muscle gene expression (25, 26) or in serum response (27, 28), two M-CAT elements described into the Troponin T promoter (29) and a CCAAT consensus sequences. More extensive analyses of the UER region demonstrate that CArG2 box and M-CAT R1 motifs are important for enhancer function. Our results show that the serum responsible factor (SRF) (27) does not bind to CArG boxes of the UER region.

## MATERIALS AND METHODS

**Cell lines and cell culture:** The mouse C2C12 and rat L6α1 skeletal muscle cell lines were maintained in growth medium (10% fetal calf serum in Dulbecco's modified Eagle's medium (DMEM)) at 37°C. For differentiation, cells were switched to differentiation medium (DM) containing DMEM with 2% fetal calf serum (FCS). Murine fibroblasts C3H10T1/2 cells were fed in growth medium containing DMEM with 15%FCS.

**Plasmids constructions, transfections and CAT assays:** 5' flanking regions of the rat *c-mos* gene were cloned upstream the chloramphenicol acetyl transferase (CAT) reporter plasmid pGEMCAT-A or pGEMCAT-C (22). The pHPmyc CAT and p5Pmyc CAT plasmids were gifts of O. Brison. Vectors 208pHPmyc CAT and 208p5Pmyc CAT harboring the insert (nt -1684 to -1476) were cloned in positive orientation into the Hind III site of pHPmyc CAT and p5Pmyc CAT respectively after filling-in the ends. Plasmids 0.8pmCH1010 and 0.2pmCH1010 contained the fragments (nt -2283 to -1476) and (nt -1684 to -1476) respectively, were filled-in the ends and cloned in positive orientation into the Hind III site (filledwith Kleenow) of the pmCH1010 plasmid (22). Vector pmCH2636 harboring the insert (nt -1684 to +139) resulted of an internal deletion of the 5' flanking region of the rat *c-mos* genomic clone λ5A1 (19 ) and was cloned in positive orientation into the polylinker of pGEMCAT-C. Vectors En1-pmCH1010, M1-pmCH1010, M2-pmCH1010 and M3-pmCH1010 contained oligonucleotides described below, cloned in positive orientation into the HindIII/PstI sites of the pmCH1010 plasmid. All constructs were confirmed by the dideoxynucleotide sequencing method .

Transfections and CAT assays were performed as previously described (22). As internal control, 1μg of the pCH110 ( Pharmacia) expressing the β–galactosidase gene under the control of the SV40 early promoter was used. All the reactions were normalized with an equal amount of protein and to ensure that any variation was encountered, transfections were realized with differents plasmid preparations, in triplicate and results were expressed after normalization of the β galactosidase activity of each transfectant.

**Gel mobility binding assays and DNAase1 footprinting:** Nuclear extracts were produced as previously described (22). Duplex oligonucleotides corresponding to the sequences of En1, En2, M1, M2, and M3) probes with HindIII and PstI ends were cloned into the multiple cloning sites of pmCH1010 construct :

**En1**:5'-AGCTTAACCAGGCTTTATCCATTTCTGAGCTGCA-3',
**M1**: 5'-AGCTTAACCAGGCTccAggCATTTCTGAGCTGCA-3'
**M2**:5'-AGCTTAACCAGGCTTTATCCggTgCgGAGCTGCA-3'
**M3**:5'-AGCTTAACCAGGCTccAggCggTgCgGAGCTGCA-3'
**SRE m1**: 5'-CTAGACAGGATGTCCATATTAGGACATCctCGT

For binding assays, 0.5 to 1ng of the UER region, 5' end-labeled on coding strand or 0.1 to 0.2ng of duplex oligonucleotides were incubated for15 min at room temperature with nuclear extracts (≈15μg) in the presence or the absence of cold competitor DNAs. DNAase footprintings were essentially performed as described in reference (22).

## RESULTS AND DISCUSSION

**A *c-mos* upstream regulatory region confers an enhancer activity in different cell lines.** We inserted different DNA fragments of the 5' upstream sequence of the rat *c-mos* genomic clone λ5A1 (19) ranging from nt -2283 to nt -1476 in positive or negative orientation upstream the ubiquitous thymidine kinase promoter. As shown in Fig.1 construct 0.8 mos pTK sense (nt -2283 to -1476) caused a twofold increase of the thymidine kinase promoter activity in L6α1 and C2C12 myoblasts while in C3H10T1/2 cells no significative activation was observed. When this DNA fragment was inserted in negative orientation (construct 0.8 mos pTK antisense), only a marginal enhancement was found (Fig.1). To delimit this positive regulatory region, insert from plasmid 0.8 mos pTK sense, was cut by MstII enzyme and recloned to give plasmids

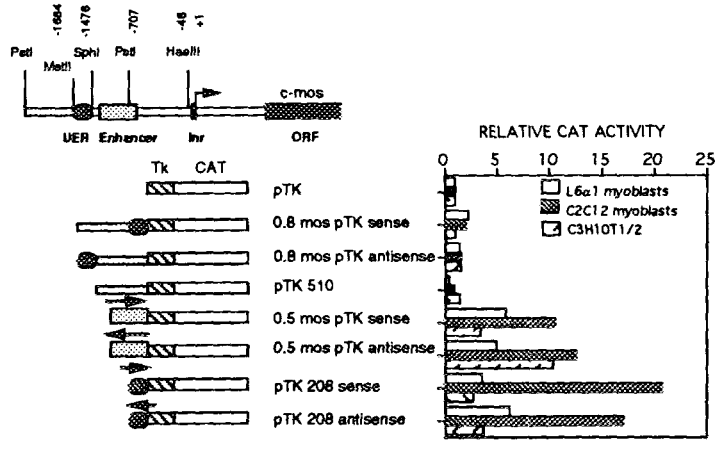


**Figure 1.** Deletion mutations of the 5' flanking regions of the rat *c-mos* gene were cloned upstream of the plasmid pTK harboring the ubiquitous thymidine kinase promoter and the CAT reporter gene.These expression vectors were transfected in transient assays into rat L6α1 and mouse C2C12 myogenic cells or into C3H10T1/2 fibroblasts. Plasmid pTK was set as 1 for each transient transfection and the fold activation of various chimeric constructs are shown. All numbers are an average of three to four transfections.

pTK510 (nt -2283 to -nt 1684), pTK208 sense and pTK208 antisense (nt -1684 to nt-1476 respectively in positive or negative orientation). No detectable activation of the TK promoter was observed with plasmid pTK510. In contrast, plasmids pTK208 sense and pTK208 antisense increased expression of the CAT activity about 4 to 10-20 fold in non-muscle and muscle cells respectively. Constructs 0.5 mos pTK sense and 0.5 mos pTK antisense containing the enhancer region described by Van der Hoorn (23) either in positive and/or in negative orientation, cis-activated 4 to10 fold the CAT activity in muscle and non-muscle cells. These results indicated that the 5' upstream region of the rat *c-mos* protooncogene between nt -1678 and nt -1470 has a high enhancer activity in somatic cell lines. In addition, the region between nt -2283 and nt -1684 (pTK510) seemed to contain a negative regulatory element which attenuated the activity of the enhancer harbored by pTK208. We named the enhancer region (between nt -1684 and -1476) UER for Upstream Enhancer Region.

**The UER acts as a strong enhancer**. We chose the promoter region of the *c-myc* gene to determine whether the UER could also affect its expression because previous report demonstrated that *c-myc* expression was under the dependence of a weak promoter (30 ). As shown in Fig.2, the minimal promoter (construct p5Pmyc; nt-2160 to nt-1620) or the 5' flanking region upstream of the *c-myc* gene (construct pHPmyc; nt-4530 to nt -1620), directed a very low level of CAT activity in C3H10T1/2 and L6α1 myoblasts. In contrast, transfections of the constructs harboring UER (208pHP myc CAT and 208p5P myc CAT) demonstrated that the UER conferred high levels of CAT activity in C3H10T1/2 and L6α1 myoblasts. These activities were higher with the minimal *c-myc* promoter (208p5P myc CAT) than with the 5' flanking region of the c-myc gene (plasmid

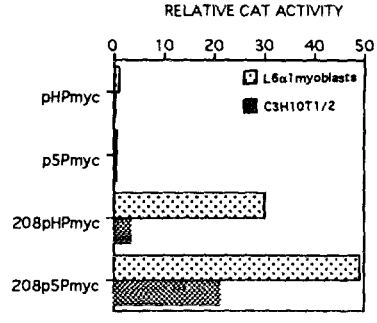


**Figure 2.** Importance of the UER element. Test constructs with the UER element driving either a part of the 5' flanking region of the *c-myc* gene or the minimal *c-myc* promoter were transiently transfected into rat L6α1 myoblasts or into C3H10T1/2 fibroblasts and results of the 208pHPmyc CAT and 208p5Pmyc CAT plasmids are given versus pHPmyc CAT and p5Pmyc CAT set as 1. All numbers are an average of four transfections.

208pHP myc CAT). The differences observed in the levels of CAT expression could be explained by the fact that the 208pHP myc CAT plasmid contains some regulatory regions interfering with *c-myc* expression (data not shown and O.Brison, personal communication). These results indicate that the UER acts as a strong enhancer in the somatic cells.

**The UER has several binding sites for nuclear proteins .** We have investigated the binding sites for the proteins involved in the interactions with the UER fragment by performing DNAase I footprinting. As shown in Fig.3A and 3B, examination of the protected sequences for known cis-acting regulatory elements reveals that the footprinted regions covered some specific DNA sequences: two CArG boxes (CArG2 in opposite orientation, nt -1570 and CArG1, nt -1520 ), 2 potential M-CAT regions, nt -1561and nt -1541 (named M-CAT-R1 and M-CAT-R2) with six of seven match for the M-CAT binding sequence already described (22) and one motifs in opposite orientation (nt -1606) including homologies with the canonical CCAAT sequences.

**Gel retardation assays show that UER does not bind SRF.** Sequence within the UER probe contains two motifs highly related to a CArG box element which binds the nuclear regulatory factor known as Serum Responsive Factor (27, 28). In order to identify whether SRF protein binds the CArG boxes of the UER and especially CArG2 element, UER were used in gel retardation assays with in vitro translated human SRF. No SRF-UER protein-DNA complex wa observed (data not shown) . However as shown in Fig.4, an oligonucleotide (SREm1) containing the wild type CArG box of the serum-responsible element of the *c-fos* promoter (31) interacted with the SRF and was efficiently competed with cold SREm1. On the other hand, the SREm1-SRF complex was not competed by oligonucleotides corresponding to UER, En1 or M1 probes. Unprogrammed lysate (RL) did not give any significant complex. These results show that SRF was unable to bind the UER element and consequently the CArG2 box and we can suppose that factor(s) that binds the CArG boxes could be a SRF related protein(s).

**CArG2 box and M-CAT R1 binding sites are important for the enhancer activity of the UER.** In order to test wether the CArG2 box and M-CAT R1 binding sites were important for the activaty of the UER, wild type and mutants of the CArG2 box and M-CAT R1 elements were inserted 5' to the pmCH1010 vector ("Inr"-CAT) reporter gene (22).Transfection of En1-pmCH1010 (CArG2 box and M-CAT R1 elements) in L6$\alpha$1myoblasts was six fold more active than the enhancerless muscle *c-mos* promoter (Fig.5).In order to find out whether CArG2 and M-CAT-R1 contributed separately or together to the enhancer activity, we tested mutants in which either CArG2 box (mutant M1) or M-CAT-R1 (mutant M2) or both CArG2 and M-CAT-R1 sites were mutated (mutant M3). CAT activity assays demonstrated that mutations within the CArG2 box (M1) and or M-CAT R1 element was sufficient to reduce significantly the transcriptional activation of these sites. Enhancer activity was completely abolished when

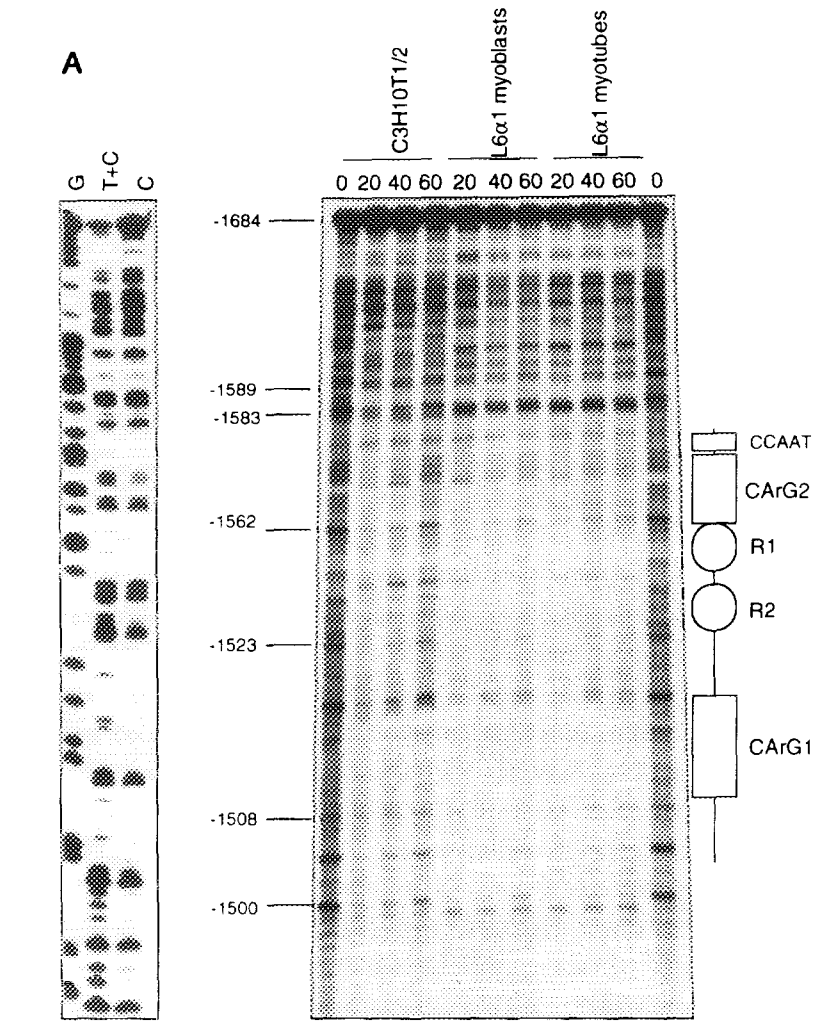


**Figure 3.** (A) DNAase I footprinting analysis of the UER. The 208bp MstII/Sph1 fragment was 5' end-labeled with $\gamma^{32}$ ATP to its coding strand. The UER was reacted with 20, 40 and 60μg of nuclear extracts from C3H10T1/2 fibroblasts, L6α1 myoblasts, and L6α1 myotubes, subjected to a DNAase I treatment and electrophoresed on a sequencing gel as described in Materials and Methods. Numbers on each side represent the nucleotide positions with regard to the start site transcription in the muscle *c-mos* initiator. (B) Sequence of the UER with consensus binding motifs identified.

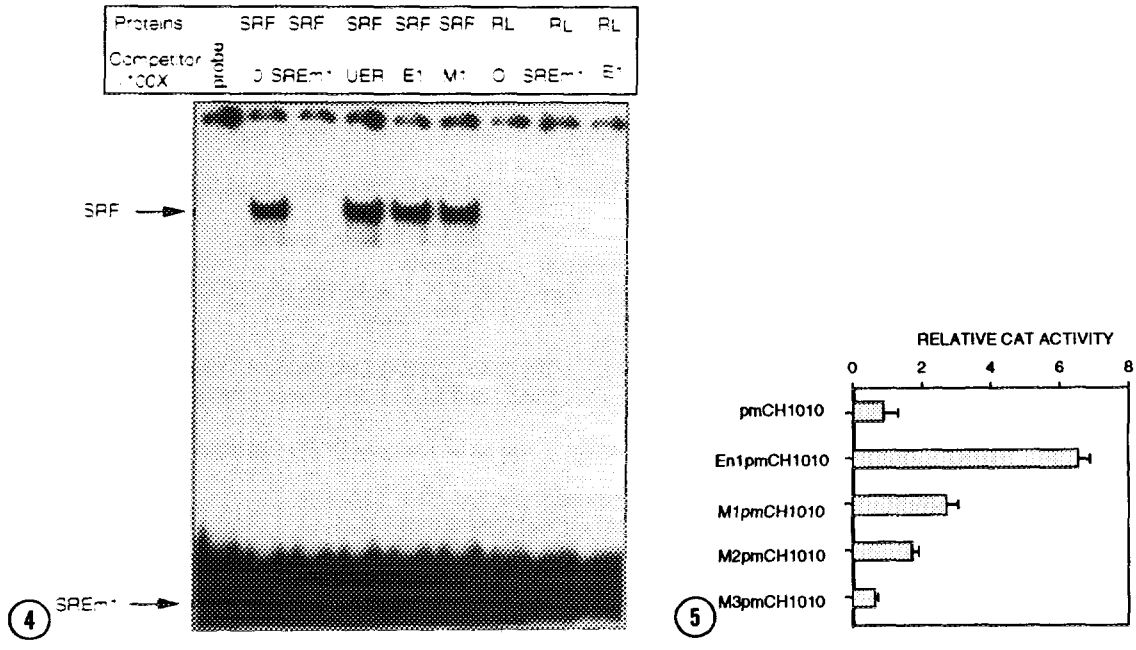


**Figure 4.** Serum Responsive factor (SRF) does not bind the CArG boxes in the UER element. An oligonucleotide containing the $CArG^{wt}$/E-$box^{mut}$ (SREm1) from the human *c-fos* SRE was used as a probe. In vitro translated SRF was incubated with the SREm1 probe in presence or not of 100-fold excesses of competitor oligonucleotides SREm1 (*c-fos* $CArG^{wt}$/E-$box^{mut}$), En1 (UER: $CArG_2^{wt}$/M-CAT-$R1^{wt}$) and M1 (UER: $CArG_2^{mut}$/M-CAT-$R1^{wt}$). Unprogrammed reticulocyte lysate (RL) was tested as control. The specific SRF-DNA complex was indicated by an arrow.

**Figure 5.** CArG2 box and M-CAT RI elements are important for enhancer activity of the *c-mos* UER. Cat activity was tested 48h after transfection of pmCH1010 fusion plasmids into rat L6α1 myoblasts. pCH110 plasmid containing the β–galactosidase gene was used as internal standard for the transfection efficiency . The specific CAT activity of each construct was calculated and normalized CAT values were divided by the normalized CAT activity of pmCH1010 (noted as value 1), resulting in the fold activation. The constructs were named, in reference to the UER element and mutated oligonucleotides as described in Materials and Methods. All numbers are an average of two to five transfection experiments.

both motifs were mutated (mutant M3). These results show that the CArG2 box and M-CAT-R1 binding sites are important elements for the transcriptional activation of the *c-mos* UER in somatic cells

In conclusion the identification of a new *c-mos* enhancer containing specific- regulatory elements permits the analysis of trans-acting factors that confer somatic patterns of regulation of this gene. Further studies will be required to identify the nature of these specific-factors and in order to test their mechanism(s) of action.

**ACKNOWLEDGMENTS:** We thank M. Guillier and K. Saintonge for technical assistance and Drs J. Harel and M.P. Leibovitch for hepful discussions. This research was supported by a grant from Association pour la Recherche contre le Cancer (A.R.C. grant n°6829).

## REFERENCES


1-Sagata, N., M. Oskarsson, T. Copeland, J. Brumbaugh, and G.F. Vande Woude.(1988) Nature 335: 519-525.

2- Sagata, N., N. Watanabe, G.F. Vande Woude, and Y. Ikawa.(1989) Nature 342: 512-518.

3- O'Keefe, S.J., A.A. Kiessling, and G.M.Cooper. (1991) Proc.Natl.Acad.Sci. USA 88: 7869-7872.

4- Paules, R.S., R. Buccione, R.C. Moschel, G.F. Vande Woude, and J.J. Eppig.(1989) Proc.Natl.Acad.Sci. USA 86: 5395-5399.

5- Yew, N., M. Oskarsson, I. Daar, D.G. Blair, and G.F. Vande Woude. (1991) Mol.Cell.Biol. 11: 604-610.

6- Hunt, T. (1989.) Nature 342: 483-484.

7- Herzog, N.K., L.S. Ramagli, and R.B. Arlinghaus. (1989) Oncogene 4: 1307-1315.

8- Leibovitch, S.A., M. Guillier, J.L. Lenormand, and M.P. Leibovitch. (1991). Oncogene 6: 1617-1622.

9- Li, C-C., H. Li, E. Chen, C.D. O'Connell, and D.L. Longo.(1993) Oncogene. 8:1685-1691.

10- Propst, F., and G.F. Vande Woude.(1985) Nature 315: 516-518.

11- Papkoff, J., I. Verma, and T. Hunter.(1982) Cell 29: 417-426.

12- Okazaki, K., M. Nishizawa, N. Furuno, H. Yasuda, and N. Sagata. (1992) Embo J. 11: 2447-2456.

13- Blair, D.G., M. Oskarsson, A. Seth, K.J. Dunn, M. Dean, M. Zweig, M.A. Tainsky, and G.F. Vande Woude.(1986) Cell 46: 785-794.

14- Daar, I., Paules, R.,S. and Vande Woude, G.,F. (1991) *J.* Cell. Biol. 114: 329-335.

15- Hannink, M., and D.J. Donoghue.(1985) Proc.Natl.Acad.Sci. USA 82: 7894-7898.

16- Leibovitch, S.A., M. Guillier, J.L. Lenormand, and M.P. Leibovitch. (1993) Oncogene 8:2361-2369.

17- Zhou, R., I. Daar, D.K. Ferris, G. Withe, R.S. Paules, and G.F. Vande Woude.( 1992) Mol.Cell.Biol. 12: 3583-3589.

18- Bai, W., B. Singh, Y. Yang, and R.B. Arlinghaus. (1992). Oncogene 7: 1757-1763.

19- Leibovitch, S.A., J.L. Lenormand, M.P. Leibovitch, M. Guillier, L. Mallard, and J. Harel.(1990) Oncogene 5: 1149-1157.

20- Van der Hoorn, F.A.(1992) Oncogene 7: 1093-1099.

21- Pal, S.K., S.S. Zinkel, A.A. Kiesling, and G.M. Cooper.(1991) Mol.Cell.Biol. 11: 5190-5196.

22- Lenormand, J.L., M. Guillier, and S.A. Leibovitch.(1993) Nucleic Acids Res. 21: 695-702.

23- Van der Hoorn, F.A.(1987) J.Mol.Biol. 193: 255-266.

24- Van der Hoorn, F.A., and S.K. Nordeen .(1989). Biochem.Biophys.Res.Commun. 161: 8-15.

25- Minty, A., and L. Kedes.(1986) Mol.Cell.Biol. 6: 2125-2136.

26- Tuil, D., N. Clergue, D. Montarras, C. Pinset, A. Kahn, and F. Phan-Dinh-Tuy. (1990) . J.Mol.Biol. 213: 677-686.

27- Treisman, R.(1986) Cell 46: 567-574.

28- Mohun, T.J., N. Garrett, and R. Treisman.(1987) Embo J. 6: 667-673.

29- Mar, J.H., and C.P. Ordahl.(1988).Proc.Natl.Acad.Sci.USA 85: 6404-6408.

30- Nepveu, A., and K.B. Marcu.(1986) Embo J. 5: 2859-2865.

31- Trouche, D., M. Grigoriev, J.L. Lenormand, P. Robin, S.A. Leibovitch, P. Sassone-Corsi and A. Harel-Bellan.(1993). Nature, 363, 79-82.