THE REGULATORY CAPACITY OF THE PROTOONCOGENE C-MYC

Thesis by

Sean Vahram Tavtigian

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1992

(Defended May 21, 1992)

Acknowledgments

I would like to acknowledge both those who have been particularly influential in my education and those who have aided in the pursuit of this thesis work.

My parents have chosen to completely support their children's education to completion. Without that support, I could not have reached this level. In junior high at Maumee Valley Country Day School, Chuck Sprandel provided encouragement and support that will never be forgotten. At Pomona College, my cooperative/competitive relationship with Cindy Scharfen (Okada) provided constant impetus towards academic achievement. The two summers spent working for James McCorquodale at the Medical College of Ohio left a lasting mark on my approach to experimental design. AT UCSF, my experiences in Keith Yamamoto's lab helped me break through the naivete fostered by my sheltered education to that point.

At Caltech, members of the Prufrock House eating group (especially Paul Garrity and Elaine Gese) and members of the Los Feliz Flyers Running and Adventure Club (especially Gary Greenberg, Jack Baldwin, and Azita Mortellero) provided fun, games, and a place to bitch and whine when I didn't have anything more constructive to say.

In this work, Barbara's introduction of conditional cell cycle progression experiments provided an addition to the logic structure of the project without which it would have been abandoned. Mike Palazzolo and Bruce Hamilton provided valuable technical advice for the cloning effort. Julianne Rogers' help with clone manipulation and management is also deeply appreciated. Finally, Barbara Wold suggested the whole thing as a rotation project in Summer, 1985. Some rotation project!

Abstract

Myc is a crucial regulatory gene capable of altering normal cellular proliferation and differentiation. Its mechanism of action is largely unknown. The cloning and identification of downstream myc regulatory targets constitutes a key step towards deeper understanding.

Beginning with a conditional c-myc expression system and a physiologic setting where conditional myc expression produced clear phenotypic effects at both cell cycle progression and mRNA expression levels, I cloned a set of myc regulated genes. The frequency with which myc targets were identified among a panel of cDNAs subject to either up or downregulation during G_1 of the first cell cycle following serum stimulated emergence from growth arrest suggested that only about one-third of such genes may be myc targets. Consequently, this work has extended the myc target gene class to include several extracellular matrix proteins, one anabolic and one catabolic enzyme, a differentiation marker, several important cell proliferation regulators, and an assortment of unidentified genes.

While that cloning effort was in progress, two groups identified max, a gene encoding multiple bHLHzip proteins that can form DNA binding oligomers either alone or in a heterotypic complex with myc. After incorporating conditional max expression into the experimental paradigm of conditional myc expression, reexamination of the myc target gene set identified individual members that are cooperatively upregulated by myc and max and members that are regulated in opposite directions by myc and by max. In addition, we made the entirely unexpected observation that max is a regulator of a specific subset of the immediate early serum response gene class.

Based on the results of these studies, I propose an integrative model accounting for the diverse effects of myc and max on cellular function.

Table of Contents

Figures and Tables

Introduction

In its most basic form, the eukaryotic cell cycle consists of alternating S- and M-phases. That simple oscillation is appropriate for cleavage stage embryos where the total mass of the embryo does not substantially increase; however, such a simple cycle provides no opportunity for growth. Addition of G_1 and G_2 to that basic cycle provides pre and post S-phase opportunities for cell growth (Keeton 1980; Cross et al., 1989).

At a nontrivial level, progression across the cell cycle phase boundaries is controlled by biochemical limitations imposed by the presence and relative activity of master regulators of cell cycle progression. The G_2/M boundary in cellular blastoderm stage Drosophila embryos provides a particularly clean example. Degradation of maternal mRNA encoding the Drosophila *cdc25* homolog *string* occurs during cell cycle 14. From that point, *string* mRNA and protein expression occur in discrete pulses just before the end of G_2. In the regulatory milieu of those cells, conditional expression of transgene *string* is sufficient to drive progression across the G_2/M boundary (Edgar and O'Farrell 1990). *cdc25* family members are protein phosphatases responsible for activation of G2 cyclin:*cdc2* family MPF (maturation or mitosis promoting factor) complexes (Dunphy and Kumagai 1991; Gautier et al., 1991; Millar et al., 1991). Frog embryos afford a biochemically homologous system where the limitation is at MPF, one step downstream of the *cdc25* homolog. Cell cycle progression in cleavage stage frog embryos can be arrested in G2 by inhibiting protein synthesis. Injection of MPF into cycloheximide treated frog embryos promotes progression across the G2/M boundary (Newport and Kirschner 1984), providing another example of a single limiting factor controlling progression across a phase boundary.

At the biochemical level, progression across the G_1/S boundary is less well understood. However, genetic analyses clearly suggest that the process is controlled in part by G1 cyclins in analogous complexes with *cdc2* homologs (Fang and Newport 1991; Koff et al., 1991; Lew et al., 1991).

In vertebrates, the cells of many adult tissues retain proliferative potential, but are arrested in a postmitotic state termed G_0 (for review, see Pardee 1989). In the better studied fibroblast tissue culture systems, overall RNA and protein synthesis rates in G_0 arrested cells are reduced, though a few transcript are upregulated following growth arrest, e.g., α-I type 3 collagen (chapter 2). Transcription factors that are members of the immediate early serum response gene class are generally present at very low levels (Lau and Nathans 1987; Almendral et al., 1988), and other transcription factors, including c-jun homodimer forms of AP-1 and the serum response transcription factor, are present in relatively inactive forms.

In G_0 arrested 3T3 cells, conditional expression of myc family (proto)oncogenes is sufficient to drive progression across the G_0/G_1 boundary, through G_1, and into S-phase (Armelin et al., 1984; Cavalieri and Goldfarb 1987; Cavalieri and Goldfarb 1988; Eilers et al., 1991). In the sense of promoting progression across a phase boundary by removing the limitation imposed by a single biochemical entity, this result is analogous to *string* and MPF promoted progression across the G_2/M boundary. However, there is also an important difference. Conditional *string* expression and MPF injection both provide the limiting activity in the regulatory milieu in which they normally function. On the other hand, conditional myc expression in G_0 arrested cells bypasses the cascade of regulatory events normally following serum or growth factor stimulation of cell surface growth factor receptors (for review, see Cantley et al., 1991). Thus the conditionally expressed protein must function in the absence of other regulators normally provided by the immediate early serum response (Lau and Nathans 1987; Almendral et al., 1988; Mohn et al., 1991) (except those that lie downstream of it) in the protein phosphatase/kinase balance of the quiescent cell.

c-myc is a member of the immediate early serum response gene class (for review, see Bravo 1990). Myc proteins are also members of the

bHLHzip transcription factor family (for reviews, see Cole 1990; Luscher and Eisenman 1990). However, the details of myc family function have proven particularly enigmatic. Homologs have not been found among the protostomes, so there has been little genetic analysis. Meaningful structural analysis followed recognition of their similarity with the bHLH and bzip families of transcription factors (for reviews, see Cole 1991; Prendergast and Ziff 1992). Those same structural homologies led to the cloning of max, a gene encoding four bHLHzip proteins that form heterodimers with myc family members (Blackwood and Eisenman 1990; Prendergast et al., 1991; Makela et al., 1992).

In this work we identify a spectrum of genes regulated by both myc and max, extending the myc target gene class to include several extracellular matrix proteins, one anabolic and one catabolic enzyme, a differentiation marker, several important cell proliferation regulators, and an assortment of unidentified genes. The experimental paradigm used in their cloning and identification should greatly facilitate expansion of the known members of their target gene class. In addition, this work in combination with reverse genetic and in vitro study of the myc:max transcription factors will finally allow direct analysis of their regulatory function.

References

Almendral, J. M., Sommer, D., MacDonald-Bravo, H., Burckhardt, J., Perera, J., and Bravo, R. (1988). Complexity of the early genetic response to growth factors in mouse fibroblasts. Molecular and Cellular Biology **8**, 2140-2148.

Armelin, H. A. M., Armelin, C. S., Kelly, K., Stewart, T., Leder, P., Cochran, B. H., and Stiles, C. D. (1984). Functional role for c-myc in mitogenic response to platelet-derived growth factor. Nature **310**, 655-660.

Blackwood, E. M., and Eisenman, R. N. (1990). Max: A helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with myc. Science **251**, 1211-1217.

Bravo, R. (1990). Growth factor-responsive genes in fibroblasts. Cell growth and differentiation **1**, 305-309.

Cantley, L. C., Auger, K. R., Carpenter, C., Duckworth, B., Graziani, A., Kapeller, R., and Soltoff, S. (1991). Oncogenes and signal transduction. Cell **64**(2), 281-302.

Cavalieri, F., and Goldfarb, M. (1987). Growth-factor-deprived BALB/c 3T3 murine fibroblasts can enter the S phase after induction of c-myc gene expression. Molecular and cellular biology **7**, 3554-3560.

Cavalieri, F., and Goldfarb, M. (1988). N-myc proto-oncogene expression can induce DNA replication in balb/c 3T3 fibroblasts. Oncogene **2**, 289-291.

Cole, M. D. (1990). The myb and myc nuclear oncogenes as transcriptional activators. Current Opinion in Cell Biology **2**, 502-508.

Cole, M. D. (1991). Myc meets its max. Cell **65**(5), 715-716.

Cross, F., Roberts, J., and Weintraub, H. (1989). Simple and complex cell cycles. Annual Review of Cell Biology **5**, 341-396.

Dunphy, W. G., and Kumagai, A. (1991). The cdc25 protein contains an intrinsic phosphatase-activity. Cell **67**(1), 189-196.

Edgar, B. A., and O'Farrell, P. H. (1990). The three postblastoderm cell cycles of Drosophila embryogenesis are regulated in G2 by string. Cell **62**, 469-480.

Eilers, M., Schirm, S., and Bishop, J. M. (1991). The myc protein activates transcription of the alpha-prothymosin gene. European Molecular Biology Organization Journal **10**(1), 133-141.

Fang, F., and Newport, J. W. (1991). Evidence that the G1-S and G2-M transitions are controlled by different cdc2 proteins in higher eukaryotes. Cell **66**(4), 731-742.

Gautier, J., Solomon, M. J., Booher, R. N., Bazan, J. F., and Kirschner, M. W. (1991). cdc25 is a specific tyrosine phosphatase that dirctly activated p34cdc2. Cell **67**(1), 197-211.

Keeton, W. T. (1980). Biological Science. New York, W W Norton.

Koff, A., Cross, F., Fisher, A., Schumacher, J., Leguellec, K., Phillipe, M., and Roberts, J. M. (1991). Human cyclin E, and new cyclin that interacts with two members of the cdc2 family. cell **66**(6), 1217-1228.

Lau, L. F., and Nathans, D. (1987). Expression of a set of growth-related immediate early genes in Balb/c 3T3 cells: coordinate regulation with c-fos or c-myc. Proceedings of the National Academy of Sciences USA **84**, 1182-1186.

Lew, D. J., Dulic, V., and Reed, S. I. (1991). Isolation of three novel human cyclins by rescue of G1 cyclin (cln) function in yeast. Cell **66**(6), 1197-1206.

Luscher, B., and Eisenman, R. N. (1990). New light on myc and myb. Part I. Myc. Genes and Development **4**, 2025-2035.

Makela, T. P., Koskinen, P. J., Vastrik, I., and Alitalo, K. (1992). Alternative forms of max as enhancers or suppressors of myc-ras cotransformation. Science **256**, 373-376.

Millar, J. B. A., McGowan, C. H., Lenaers, G., Jones, R., and Russel, P. (1991). p80cdc25 mitotic inducer is the tyrosine phosphatase that activates p34cdc2 kinase in fission yeast. European Molecular Biology Organization Journal **10**(13), 4301-4309.

Mohn, K. L., Laz, T. M., Hsu, J.-C., Melby, A. E., Bravo, R., and Taub, R. (1991). The immediate-early growth response in regenerating liver and insulin-stimulated H-35 cells: comparison with serum stimulated 3T3 cells and identification of 41 novel immediate-early genes. Molecular and Cellular Biology **11**(1), 381-390.

Newport, J. W., and Kirschner, M. W. (1984). Regulation of the cell cycle during early Xenopus development. Cell **37**, 731-742.

Pardee, A. B. (1989). G1 events and regulation of cell proliferation. Science **246**, 603-608.

Prendergast, G. C., Lawe, D., and Ziff, E. B. (1991). Association of Myn, the murine homolog of max, with c-myc stimulates methylation-sensitive DNA-binding and ras cotransformation. Cell **65**(3), 395-407.

Prendergast, G. C., and Ziff, E. B. (1992). A new bind for myc. Trends in Genetics **8**(3), 91-96.

Chapter 1

Conditional expression of c-myc

Sean V. Tavtigian and Barbara J. Wold

Introduction

The diverse assays in which effects of the c-myc oncoprotein are observed suggest it plays significant roles regulating cell proliferation and differentiation. During embryonic development, c-myc is strongly expressed in proliferating endodermal and mesodermal tissues (Zimmerman et al., 1986; Schmid et al., 1989). Overexpression and deregulation of c-myc are often observed in naturally arising neoplasms, and its ectopic expression in transgenic mice vastly increases the incidence of a variety of tumors (Leder et al., 1986). Expressed in tissue culture cells, c-myc exerts an effect on growth factor requirements, can drive growth arrested cells into S-phase, antagonizes cellular differentiation in some settings, and has been reported to modulate gene activity at both transcriptional and post transcriptional levels (Prendergast and Cole 1989; for reviews, see Bishop 1986; Cole 1986).

The network of biochemical pathways leading from growth factor stimulation of a quiescent cell to completion of its cell cycle are clearly complex and interactive; present data indicate that several distinct signalling pathways proceed via steps involving myc proteins (Armelin et al., 1984; Rapp et al., 1985; Cavalieri and Goldfarb 1988). Proliferating fibroblasts synthesize both c-myc RNA and protein continuously (Thompson et al., 1985), but their abundance decreases dramatically when the cells are arrested in G_0. Stimulation of growth arrested cells with serum or platelet derived growth factor (PDGF) induces c-myc accumulation concomitant with re-entry into the cell cycle (Persson et al., 1985; Thompson et al., 1985). A direct role for myc is shown when its forced expression in growth arrested cells induces re-entry into the cell cycle (Armelin et al., 1984; Cavalieri and Goldfarb 1987). Furthermore, blockage of myc protein synthesis during serum stimulation with antisense oligonucleotides apparently antagonizes progression to S-phase (Heikkila et al., 1987).

C-myc has long been postulated to be an important regulator of gene activity. Its product is a nuclear phosphoprotein, and the timing of its expression following serum stimulation of quiescent cells suggests that it acts before initiation of DNA synthesis. The primary structure of c-myc shares with the adenovirus E1a proteins and SV40 large T antigen two copies of a structural motif (Figge et al., 1988), required in ras cotransformation assays, which may be either a retinoblastoma protein binding site and/or a phosphorylation site (Dyson et al., 1989; Ewen et al., 1989; Luscher et al., 1989). Near its C-terminus the protein contains a basic region adjacent to a helix loop helix domain, a motif mediating multimerization and DNA binding in a number of differentiation effector proteins (Caudy et al., 1988; Murre et al., 1989; Wright et al., 1989). Contiguous with the second loop of the helix loop helix domain is a leucine repeat, another structure mediating protein multimerization. Among the proteins sharing regions of similarity with c-myc are known transcriptional regulators; current data indicate that many are sequence specific DNA binding proteins; (Myers et al., 1981; Hauschka and Weintraub, unpublished).[1] Others, including the Adenovirus E1a proteins, appear not to require DNA binding to exert their regulatory activities (Hoeffler et al., 1988; Reichel et al., 1988).

Here we have established cell lines stably transfected with a metal inducible c-myc gene. Conditions were defined under which myc RNA and protein levels were effectively regulated. We report the effect of transient and prolonged myc induction on the cell cycle and on the activity of a series of genes, including several previously reported to be targets of c-myc regulation.

---

[1] Basic helix loop helix domain differentiation effector proteins appear to function as heterodimers (Lassar et al., 1991), but the possibility remains that some structurally related transcription factors such as USF/MLTF and AP-4 may function as homodimers (Gregor et al., 1990; Hu et al., 1990). At very high protein concentrations c-myc will form DNA binding homodimers, but a recently cloned pairing partner of c-myc, max/myn protein (Blackwood and Eisenman 1990; Prendergast et al., 1991), preferentially forms DNA binding heterodimers with c-myc and cooperates strongly with c-myc in ras cotransformation assays.

Materials and methods

Recombinant plasmids. Plasmids were constructed essentially according to the methods of Maniatis et al. (Maniatis et al., 1982). pMyc23c (figure 1) encodes the mouse c-myc proto-oncogene and was used to stably transfect NIH 3T3 cells. The plasmid was assembled in several steps in pT7T3-18 (Bethesda Research Laboratories, Gaithersburg, MD). From 5' to 3', it consists of a 756 bp Taq I - Sac I fragment from the MTV LTR bearing a glucocorticoid regulated enhancer (Payvar et al., 1983), a 218 bp Sac I - Bgl II fragment containing the mouse metallothionein-I promoter (Mueller et al., 1988), a 550 bp Xba I - Sac II fragment from the first intron and second exon of the mouse c-myc genomic clone lambda myc 5 (Hood and Barth, unpublished), a 922 bp Sac II - Nsi I fragment containing the remainder of the c-myc coding sequence derived from the c-myc cDNA myc 29 (Kim, Mueller, and Wold, unpublished), and a 302 bp Hgi AI - Hind III fragment containing the poly-A addition signal from the mouse Mt-1 gene (Mueller et al., 1988). The Taq I site at the 5' end of the enhancer fragment was filled-in and ligated to a filled-in Eco RI site within the polylinker, regenerating the Eco RI site. The Bgl II site at the 3' end of the Mt-I promoter was ligated into the Bam HI site of the polylinker, destroying both sites; similarly, the Nsi I - Hgi AI fusion between the 3' end of the myc cDNA fragment and the 5' end of the Mt-I poly-A addition fragment destroyed both sites.

Plasmid pNeo 3, an HSV tk promoted neo construct (Bond and Wold 1987), was used as a G418 selectable marker in the preparation of stably transfected cell lines. pRSV lacZ (W. Albert, unpublished) was used as an internal control in transient cotransfection assays, as was pMSV CAT-S (Harland and Weintraub 1985). Plasmid pHHS CAT contained a 1 kb fragment of a human hsp 70 promoter (Drabent et al., 1986) which had been deleted to +26 and fused to the gene encoding chloramphenicol acetyl transferase.

The following plasmids were assembled in T7 T3 duel promoter vectors for synthesizing RNA hybridization probes. pMyc2-XS contains an approximately 940 bp Xba I - Sac I fragment spanning the entire second exon of c-myc. pMtmyc2-SR contains a 416 bp SacI - Eco RV fragment from pMyc23c including the 5' end of the transcript and extending into c-myc sequences. pMyc1-HA contains a 122 bp Hind III - Alu I fragment from the first exon of c-myc. The Hind III site is located between P1 and P2 (Bernard et al., 1983) within the c-myc promoter. pMHS-SX contains a 646 bp Sma I - Xba I internal fragment from the mouse HSP 70 cDNA MHS 214 (Lowe and Moran 1986). pFos-PP2 contains a 108 bp Pvu II - Pst I fragment spanning the 5' end of the c-fos transcript (Van Beveren et al., 1983). pp53-XA contains a 1005 bp Xho I - Asp 718 fragment from the p53 tumor antigen gene (Tan et al., 1986). pB5 contains an approximately 135 bp 3' end fragment from a mouse tubulin β5 cDNA (Lewis et al., 1985) from which the poly-A tract has been trimmed with exonuclease Bal 31. pCAT-SPII contains a 153 bp Sau IIIA - Pvu II fragment from the gene for chloramphenicol acetyl transferase in which two lysine codons (3 and 4) were changed from AAA to AAG by a site directed mutagenesis. pRSV lacZ-ES contains a 213 bp Eco R1 - Sau 3A fragment of pRSV lacZ spanning the 5' end of the RSV transcript and extending into Lac Z sequences. Unexpectedly, RNase protections of RSV lacZ transcripts with the RSV lacZ probe yield two protected fragments. One protection product was consistent with a transcript initiating at the RSV promoter. The second protected fragment was shorter; its structure has not been investigated

Cell culture. NIH 3T3 cells were maintained in Dulbecco modified Eagle medium (DME) supplemented with 10% calf serum (Gibco) or 10% fetal calf serum, (Hyclone Laboratories and Hazleton Biologics), as specified in individual experiments, 100 U of penicillin G per ml, and 7.5 U of streptomycin per ml. Cells were transfected with a calcium phosphate co-precipitate of plasmid and carrier DNAs by a standard procedure (Wigler et al., 1979). Using pNeo 3 as a selectable marker,

Table 1 Probe descriptions. All RNA analyses were solution hybridization RNase protection assays. The abundance of the various RNA species examined varied widely; the specific activity at which each probe species was synthesized was chosen to partially normalize the resultant signals, and was held constant from experiment to experiment

a. Units of sensitivity are: $transcripts^{1} \times 20\ cpm^{-1} \times cell^{-1}$

b. For transcripts originating at P1. Transcripts initiating upstream of P1 (for instance, at P2) will protect longer probe fragments.

c. Because of sequence differences between the MHS 214 probe and the MHS 213 and 243 messages, protection of these messages by the MHS 214 probe results in a series of bands about 85 nucleotides in length.

d. The B5 probe is included in the same hybridization with either p53-XA or RSVlacZ-ES; consequently, either 1 or 4 ug of RNA is used per hybridization, respectively. The sensitivity of the hybridization is compensated by synthesizing the probe at 4000 Ci/mMol when 1 ug of RNA is used and synthesizing it at 1000 Ci/mMol when 4 ug of RNA is used.

Probe	RNA species protected	Protected fragment length, nucleotides	Specific activity of protected fragment, curies per mMole	ug of total RNA per hybridization	sensitivity[a]
MTmyc2-SR	MTmyc23C transgene	257	5,000	1.0	60.2
	c-myc, exon 2	159	3,400		88.5
	metallothionein-1	66	850		354.1
Myc1-HA	c-myc, exon 1	99.[b]	20,000	2.0	7.5
MHS-SX	hsp70 MHS 214	130	40,000	2.0	3.8
	hsp70s MHS 213 / 245	85.[c]	20,000		7.5
Fos-PP2	c-fos	96	20,000	1.0	15.1
p53-XA	p53 tumor antigen	175	20,000	1.0	15.1
B5	tubulin B5	129	4,000	1.0[d]	75.3
CAT-SP	transfected MSV-CAT transcript	156	20,000	4.0	3.8
	transfected HHS-CAT transcript	145	18,000		4.2
RSVlacZ-ES	transfected RSV-lac Z transcript	220	30,000	4.0	2.5

NIH 3T3 cells co-transfected with pMyc23c were initially selected for growth in 600 µg/ml G418 (Gibco). After 3 days, the G418 concentration was reduced to 400 µg/ml, and individual clones were maintained in 200 µg/ml G418. In the first two sets of transfections, relatively fast growing clones were chosen for individual analysis, while remaining clones were pooled. Of more than fifteen individual clones checked, none synthesized more RNA from the c-myc transgene than from their endogenous c-myc gene. In contrast, the pooled cell lines expressed considerably more myc RNA. In a third transfection, slow growing clones were deliberately selected. Of 10 clones tested, 4 expressed high levels of myc23c RNA. Initial analysis of RNA from these cells using probes sensitive to the 5' and 3' ends of the Mtmyc23c transgene showed that the predicted myc23c transcript was made. Subsequently, several individual myc23c expressing lines were reisolated from the myc23c expressing pools. Although the cell lines were initially selected and propagated in DME plus 10% fetal bovine serum (Hyclone Laboratories), they were transferred to DME plus 10% calf serum (Gibco) for routine cell culture through the subsequent experiments.

Culture in defined medium was essentially as in Zhan and Goldfarb (Zhan and Goldfarb 1986), with the modifications noted below. Tissue culture dishes (6 cm, 10 cm, and 24 well, Nunc) were first coated with 1 mg/ml polylysine for 2-3 hours at 37°, rinsed with phosphate buffered saline (PBSA) , and coated overnight at room temperature with 15 µg/ml bovine fibronectin in phosphate buffered saline. Defined medium (DMI) consisted of a 3:1 mixture of DME and Ham's F12 (Flow Laboratories, McLean, VA) supplemented with 15 mM HEPES pH 7.4, 4 µM manganese chloride, 10 µM ethanolamine, 100 nM sodium selenite, 5 µg/ml transferrin, and 500 µg/ml bovine serum albumin-linoleic acid complex. We did not add histidine or hydrocortisone. Hormonal supplements were 30 ng/ml insulin and 10 pg/ml transforming growth factor B (RD Systems, Inc.). We found that low concentrations (10 pg/ml is $1/10^{th}$ the ED_{50} for stimulation of soft agar colony formation by NRK fibroblasts

(Roberts et al., 1985) of TGF-β, which should facilitate assembly of exogenous fibronectin (Allen-Hoffman et al., 1988), improved cell survival through the timecourse of the experiments.

RNA preparation and measurement. For zinc titrations in serum containing media, 10 cm plates were seeded with $4x10^5$ cells in 10 ml of medium. Forty eight hours later, DME or DME containing $ZnSO_4$ was added to the concentrations indicated in the figure legend. Volumes added ranged from 0.45 to 1.0 ml. Early in this work we noticed that control inductions performed by removing the medium and replacing it with either new or conditioned medium resulted in transient induction of genes including c-myc and metallothionein-I. This problem was largely eliminated by adding the inducing agents in 5-10% of the volume of the medium in the culture dish. Cells were harvested 4 hours later. For zinc titrations in defined media, 10 cm polylysine/fibronectin coated plates were seeded with $1x10^6$ cells. Twelve hours later, cells were washed once with PBSA and fed with 10 ml of DMI. Twelve hours later, cells were refed with DMI. Twenty four hours later, 3:1 DME:F12 with or without zinc was added in the same manner as above.

For timecourse inductions, 10 cm polylysine/fibronectin coated plates were seeded with 3T3 cells at $7.5x10^5$, or with the myc overproducing lines 10-2 or 3-5B at $1.2x10^6$ per plate. Twelve hours later cells were cotransfected with a $CaPO_4$ co-precipitated mixture containing 1 μg of pMSV CAT-S, 3 μg of pHHS CAT, 6 μg of pRSV lacZ, and 15 μg of calf thymus DNA. Cells were 10% DMSO shocked 10 hours later, then fed with 10 ml of DMI. Cells were refed with DMI 12 hours later. Twenty four hours later, cells were induced with the addition of 0.5 ml of fetal calf serum or 1 ml of 3:1 DME:F12 with or without 220 μM $ZnSO_4$. Cells were harvested according to the timecourse indicated in the figure legends.

RNA was prepared by a modified guanidinium thiocyanate-organic extraction protocol. The guanidinium solution was 4M guanidinium thiocyanate, 50 mM Tris pH 7.6, 12.5 mM EDTA, and 2% N-lauryl sarcosine.

Aliquots of this solution were adjusted to 1% 2-mercaptoethanol just before use. To harvest RNA from a 10 cm culture dish, the cells were lysed on the plate in 1 ml of guanidinium solution. The lysate was transferred to a 15 ml polypropylene tube and incubated at 60° for 5 minutes. 3 ml of prewarmed phenol, 3 ml of chloroform, and 3 ml of prewarmed 100 mM sodium acetate pH 4.5 were added sequentially at 5 minute intervals and shaken thoroughly. After addition of the sodium acetate the RNA prep was incubated at 60° for 10 minutes and then at 0° for 10 minutes. Samples were centrifuged at 12,000g for 10 minutes at 4°. The aqueous phase was reextracted with ether. RNA was precipitated by the addition of 1 ml of 10 M ammonium acetate and 5.5 ml of isopropanol. The pellet was redissolved in TE (10 mM Tris pH 7.5, 1 mM EDTA), transferred to a 1.5 ml eppendorf tube, adjusted to 300 mM sodium acetate, ethanol precipitated, and redissolved in TE. We observed that RNA prepared from transiently transfected cells by this method often contain enough residual plasmid DNA to result in a significant hybridization background. Digestion of portions of the RNA prep with 1 unit of DNase I per µg of nucleic acid eliminated this background.

RNA probes were synthesized essentially as recommended by the enzyme manufacturer. Full length probes were excised from 5% acrylamide gels crosslinked with BAC (N,N'-bis-acrylylcystamine, Bio-Rad) and the acrylamide was dissolved with 200 mM 2-mercaptoethanol in TE. RNase protections were performed by a modification of the procedure described by Zinn et. al. (Zinn et al., 1983). One to 4 µg of total cellular RNA plus tRNA to make a total of 25 µg were coprecipitated with probe. Samples were redissolved in 30 µl of RNA hybridization buffer,(40 mM Pipes pH 6.7, 400 mM NaCl, 1 mM EDTA, 80% formamide) denatured at 80° for 10 minutes, and hybridized for >12 hours. Probe concentrations and hybridization temperatures were optimized so that hybridizations proceeded to at least 90% of completion in 12 hours. Hybridizations were digested with the addition of 300 µl of 1000 Units/ml RNase T1 (US Biochemicals) in RNase digestion buffer (10 mM Tris pH 7.5, 300 mM NaCl,

5mM EDTA) for 60 minutes at 30°, followed by 200 µl of 250 µg/ml proteinase K, 1% SDS for 15 minutes at 30°. Samples were prepared for electrophoresis, fractionated on 5% denaturing acrylamide gels, dried, and exposed to X-ray film. Handling of samples containing dissolved BAC acrylamide was facilitated by co-precipitating RNA samples and probe out of 100 mM 2-mercaptoethanol, adjusting gel loading buffer to 1% 2-mercaptoethanol, and using well silanized tubes.

Immunocytochemical staining. Cells were grown in Leighton tubes (Costar). Adherent cells were rinsed with PBS and fixed for 10 minutes in buffered 10% formalin (American Scientific Products). The cells were rinsed with PBS and sequentially immersed in: PBSA, 10 minutes; -20° acetone, 7 minutes; and PBSA, 10 minutes.

Immunoperoxidase staining employed Vectastain ABC kit reagents diluted according to the manufacturer's (Vector Laboratories) directions. The cells were covered with diluted blocking serum for 20 minutes prior to 16-20 hour, 37° primary antibody incubations in a humidified chamber. Six different antibody solutions localized in patches were incubated on each Leighton tube insert. Anti-human c-myc antibodies were prepared against recombinant protein purified from E. coli (R. Koski and W. Bryce, unpublished data). For competition experiments, antibodies were pre-incubated with recombinant c-myc protein for 30 minutes at 37°. Biotinylated anti-mouse IgG or anti-rabbit IgG secondary antibody incubations were for 30 minutes at room temperature. The avidin-biotinylated horseradish peroxidase ABC complex was applied to the cells for 60 minutes. The 0.5 mg/ml 3,3'-diaminobenzidine tetrahydrochloride / 0.01% H_2O_2 substrate incubations were for 6 minutes. The stained cells were immersed in: H_20, 5 minutes; 95% ethanol, 3 minutes (twice); 100% ethanol, (twice); dipped 15 times in xylene; and coverslipped with Permount (Fisher).

DNA Replication Assays. Polylysine/fibronectin coated 24 well plates were seeded at a density of $2x10^4$ cells per well in 0.5 ml of DME plus 10% calf serum. Four hours later each well was washed with 1 ml of

PBS and refed with 0.75 ml of DMI. Twelve to 18 hours later the cells were refed with 0.75 ml of DMI. Twenty four hours later, 3:1 DME : F12, $ZnSO_4$ diluted in 3:1 DME : F12, growth factors, and/or fetal bovine serum was added to the wells. Volumes added ranged from 15 to 37.5 µl. For zinc and hormone titration experiments, entry into S-phase was monitored by adding 37.5 µl of 50 uCi/ml tritiated thymidine to each well 14 hours after stimulation. In timecourse induction experiments, zinc was added as above; subsequently, zinc containing media were removed at the times indicated in the legend, and the cells were washed and refed with DMI. Again, tritiated thymidine was added 14 hours after stimulation. Experiments were terminated 10 hours after thymidine addition by washing the cells with TBS (50 mM Tris pH 7.5, 0.9% NaCl) 3 times and then fixing them in 1:1 methanol : acetone for 90 seconds. After drying, the wells were coated with warmed Kodak NTB4 nuclear track emulsion and exposed for 36 hours at room temperature before being developed. Cell counts were aided by staining nuclei briefly with 1 µg/ml Hoechst 33342, an intercolating fluorescent dye which soaks through dried emulsion efficiently.

Southern Blots. DNA was prepared from cells scraped from one 10 cm plate by digesting them in 800 µl of 250 µg/ml proteinase K, 0.5% SDS, 400 mM NaCl, 2 mM EDTA, 20 mM Tris pH 7.5 overnight at 55°. The digests were phenol chloroform extracted twice, ether extracted once, and precipitated by the addition of 125 µl of 60% polyethylene glycol. The DNA was washed with 100% ethanol and 70% ethanol, dried, and redissolved in TE.

Cellular DNAs were digested with appropriate restriction enzymes and fractionated on an 0.8% agarose gel. After denaturation and neutralization, the DNA was transfered to Hybond N (Amersham) in 10 X SSPE. Filters were baked for 2 hours at 80°. Prehybridization solution was 6X SSPE, 50% formamide, 1% SDS, 1 mg/ml polyvinylpyrolidone, 200 µg/ml tRNA, 0.1% tween 20, and 10% dextran sulfate. Hybridzation solution was the same with the addition of $2x10^6$ cpm/ml of an anti-

sense RNA probe synthesized from pMyc2-XS at a CTP specific activity of 400 Ci/mMole. Prehybridization was at 50° for 1 hour, hybridization was at 50° for 12 hours, and the final wash was in 0.1X SSPE, 0.1% SDS at 68°.

Western Blots. Polylysine/fibronectin coated 6 cm plates were seeded at a density of 2.5-4.0 x 10^5 cells per plate in DME plus 10% calf serum. Twelve hours later, plates cells were washed with PBSA and fed with 3 ml of DMI. Twelve hours later cells were refed with 3 ml of DMI. Twenty four hours later cells were induced with 300 µl of 3:1 DME:F12, 150 µl of fetal bovine serum (Hazleton Biologics), or 300 µl of 220 µM $ZnSO_4$ in 3:1 DME:F12. Cells were harvested over a 16 hour timecourse as indicated in the figure legend. Cells were lysed in 200 µl of loading buffer (80 mM Tris pH 6.8, 100 mM dithiothreitol, 2% SDS, 10% glycerol). Cell lysates were fractionated on acrylamide SDS gels and electroblotted to Nitroplus 2000 (Micron Separations). Filters were blocked for 2 hours at room temperature in 1X TBS, 0.5% Tween 20, 2% BSA. Antibody incubations were in the same solution with the addition of 2 µg/ml anti-myc monoclonal antibody H60C37 ^{125}I labelled to a specific activity of approximately 100 Ci/mMole. Antibody incubations were for 12 hours, followed by five 20 minute washes in blocking solution.

Results

Preparation and characterization of cell lines. We have prepared a series of NIH 3T3 cell lines which express a stably transfected, zinc-inducible c-myc gene. The inducible c-myc gene (Mtmyc23c) (figure 1) consists of the glucocorticoid regulated enhancer from murine mammary tumor virus and the mouse metallothionein-I (Mt-I) promoter upstream of sequences encoding mouse c-myc. The 3' untranslated region of c-myc, which includes sequences specifying relatively rapid RNA turnover, (Jones and Cole 1987; Schuler and Cole 1988), was replaced with the 3' untranslated region of Mt-I. This plasmid was cotransfected with pNeo3 (Bond and Wold 1987) into NIH 3T3 cells. G418 resistant clones were screened for zinc induced expression of either transgene mRNA or overall c-myc protein levels. Two independent transformants are further characterized here.

In initial characterizations of induction conditions, the effects of $ZnSO_4$, $CdSO_4$, and dexamethasone on accumulation of the myc23c and endogenous c-myc transcripts were examined. As a control, the endogenous β5 tubulin transcript, expression of which is not expected to be affected by the inducing agents, was also monitored. Zinc and cadmium induction gave similar results, but the effective concentration of cadmium was always significantly lower than the effective concentration of zinc. Induction of the myc23c transcript by dexamethasone was tested and proved unexpectedly transient and inconsistent; the matter was not pursued further. Figure 2 presents data comparing transcripts levels in the myc23c expressing cell line 10-2 with parental NIH 3T3 cells as a function of zinc concentration. Growing exponentially in medium containing 10% calf serum (figure 2a), the cells tolerated zinc concentrations up to 100 μM. In 10-2 cells basal expression of the myc23c message was easily detected, and its abundance increased from about 10 transcripts per cell to several hundred in a zinc dependent manner. At low concentrations of zinc, 10-2

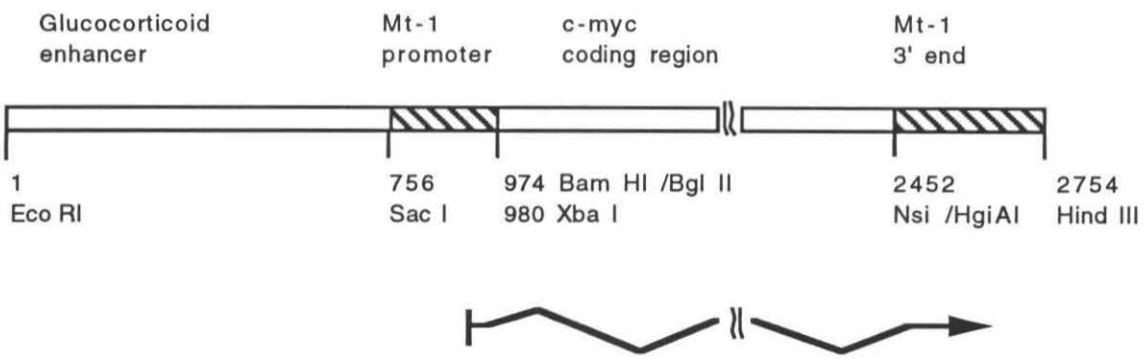
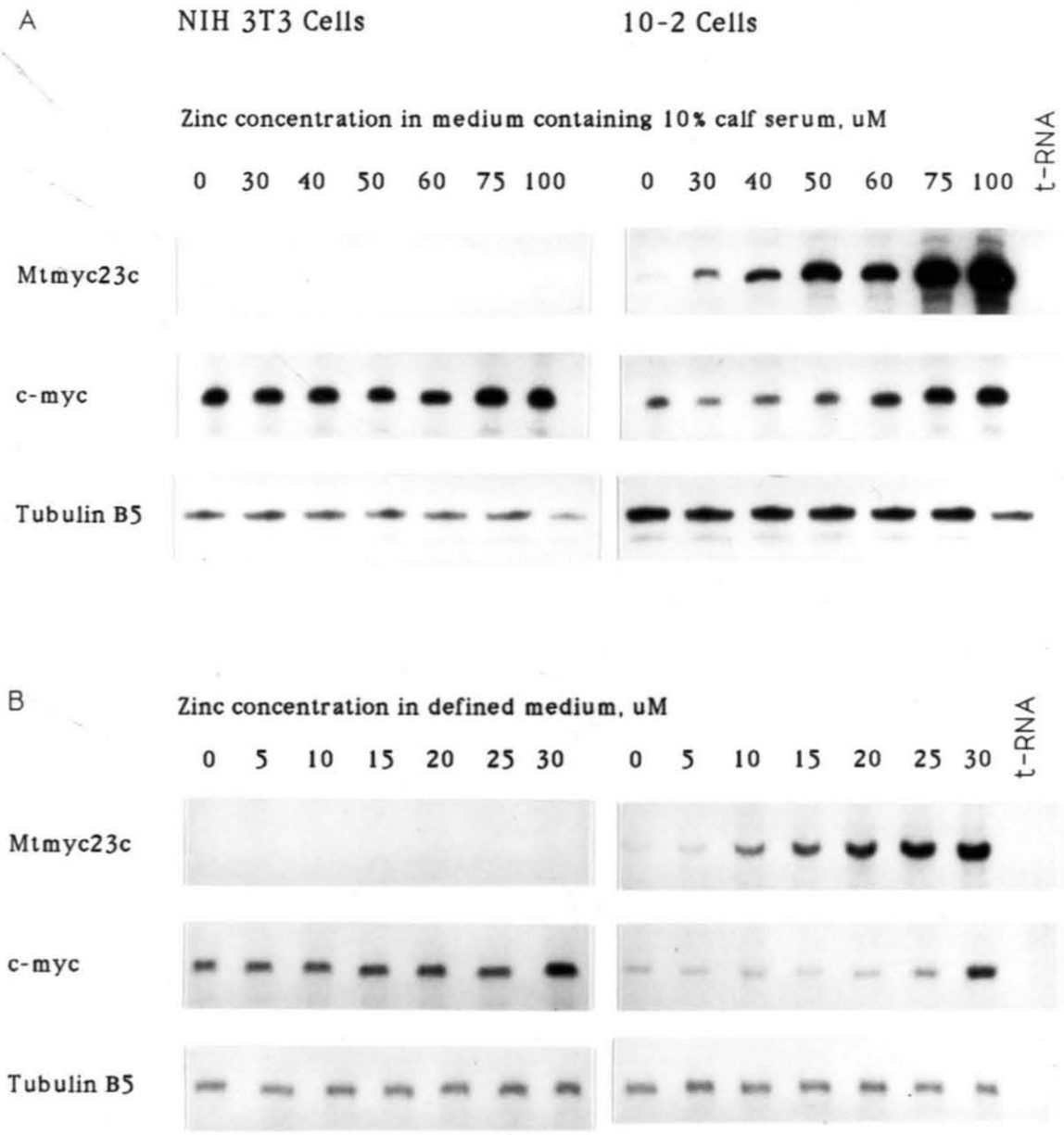


Figure 1. Plasmid myc23c encodes mouse c-myc. The structural gene has been linked to the glucocorticoid regulated enhancer from murine mammary tumor virus (Payvar et al., 1983) and the mouse Mt-I promoter (Mueller et al., 1988). The Bam HI / Bgl II and Nsi I / Hgi AI junctions are no longer recognized by the enzymes used in their creation. The wavy line denotes the expected transcript initiating within the Mt-I promoter and terminating at the Mt-I polyadenylation signal.

cells contained 2- to 4-fold less endogenous c-myc RNA than parental 3T3 cells, but their endogenous c-myc RNA levels increased with added zinc until they equaled those of 3T3 cells at high zinc concentrations, about 20 transcripts per cell. The abundance of tubulin β5 RNA (Lewis et al., 1985) was essentially invariant within each cell line. Anticipating cell cycle progression studies of growth arrested cells, a context which provides a clear phenotype from conditional c-myc overexpression (Cavalieri and Goldfarb 1987; Cavalieri and Goldfarb 1988), transcript levels were examined in cells which had been growth arrested in defined medium before treatment with zinc (Zhan and Goldfarb 1986) (figure 2b). The pattern of transcript regulation was qualitatively similar. The abundance of the myc23c transcript was 2- to 3-fold lower in either the presence or the absence of zinc inducer than in cells growing in serum containing medium, but remained highly inducible by zinc. As expected, both 3T3 and 10-2 cells expressed less endogenous c-myc transcript when they were growth arrested than they did while growing exponentially. Expression of tubulin β5 RNA was both reduced and essentially invarient across the titration.

At zinc concentrations above 10 μM in defined medium or above 40 μM in medium supplemented with 10% calf serum, the abundance of the myc23c transcript in 10-2 cells exceeded the abundance of c-myc RNA in the control cells. However, the physiological effects of c-myc expression presumably depends on c-myc protein, and we could not predict *a priori* the level of myc protein that would accumulate from the myc23c transcripts. Therefore, myc protein levels were monitored by Western blot, probing with an iodinated monoclonal antibody to c-myc. Control NIH 3T3 cells and two myc23c expressing lines, 10-2 and 3-5B, were growth arrested in defined medium; c-myc expression in 3T3 cells was stimulated with serum while c-myc expression in 10-2 and 3-5B cells was stimulated with zinc. In a pairwise comparison of lysates prepared from uninduced and induced cells (figure 3a), we found that growth arrested 3-5B cells (as assayed by ^{3}H thymidine incorporation) contain as much

Figure 2. C-myc expression as a function of zinc concentration. **A.** Zinc titration in medium containing 10% calf serum. NIH 3T3 cells and myc23c expressing 10-2 cells were seeded on 10 cm plates at a density of $4x10^5$ cells per plate. Forty eight hours later, the media were adjusted to the concentration of $ZnSO_4$ indicated at the tops of the lanes (0 - 100 μM). Cells were harvested for RNA preparation after 4 hours of induction. **B.** Zinc titration in defined medium. NIH 3T3 cells and 10-2 cells were seeded and brought to growth arrest as indicated in materials and methods. The media were then adjusted to the concentration of $ZnSO_4$ indicated at the tops of the lanes (0 - 30 μM). Cells were harvested for RNA preparation after 4 hours of induction. tRNA lane: hybridization containing yeast tRNA but no sample RNA. The analysis was by RNase protection; data on the probes used to detect the indicated messages are presented in materials and methods.


A
NIH 3T3 Cells
10-2 Cells
Zinc concentration in medium containing 10% calf serum, uM
0 30 40 50 60 75 100
0 30 40 50 60 75 100
t-RNA
Mtmyc23c
c-myc
Tubulin B5
B
Zinc concentration in defined medium, uM
0 5 10 15 20 25 30
0 5 10 15 20 25 30
t-RNA
Mtmyc23c
c-myc
Tubulin B5

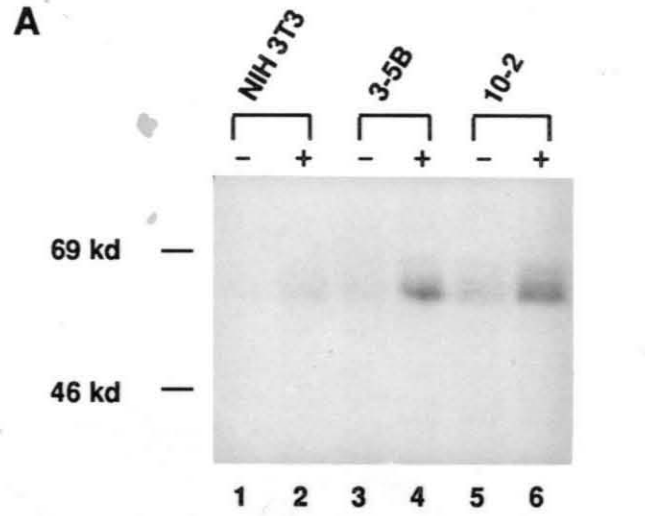
c-myc protein as serum stimulated 3T3 cells, while 10-2 cells express about twice as much. In a timecourse analysis (figure 3b), accumulation of myc protein in zinc stimulated 3-5B and 10-2 cells peaked after 2 hours of induction, but even after 16 hours of induction was still several-fold higher than steady-state uninduced levels. Immunocytochemical staining of cells with anti c-myc monoclonal and polyclonal antibodies confirmed the expected nuclear localization of c-myc protein synthesized during zinc inductions (data not shown).

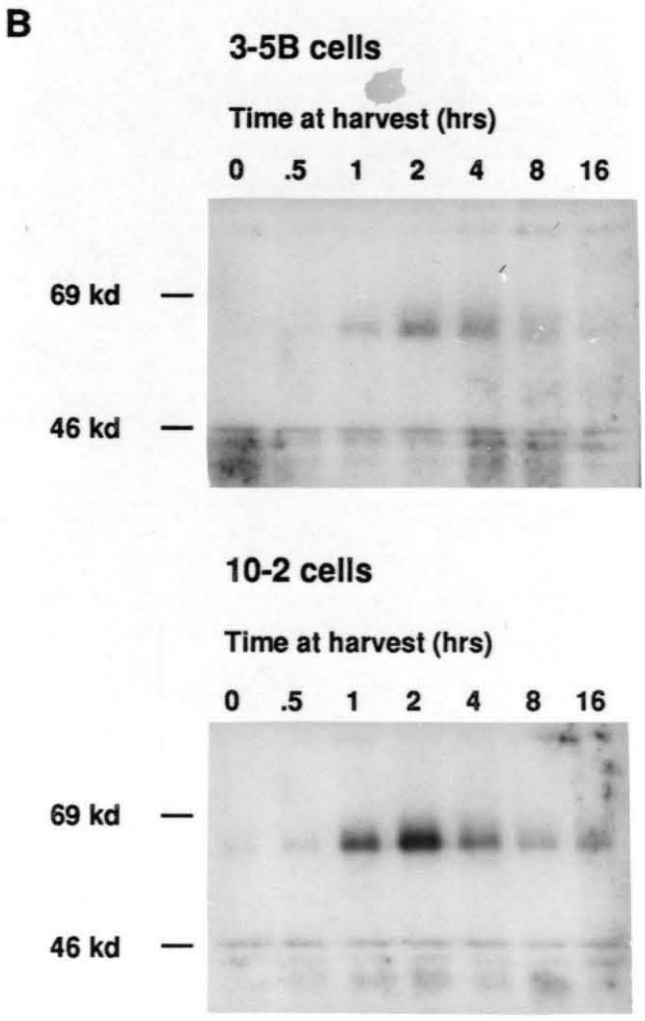
Myc-dependent cell cycle progression. We examined c-myc mediated progression of quiscent cells through G1 to S-phase. Several points were of particular interest. First, we wished to demonstrate biologic activity of overproduced c-myc in this system; cell cycle progression experiments provide an appropriate assay. Second, previous studies, in different cell hosts, have uncovered synergistic cooperation between myc overproduction and the activities of specific growth factors; we sought to define a cooperative interaction which would increase the magnitude of c-myc's regulatory activity in this system. Finally, this inducible system afforded the opportunity to examine the duration of c-myc expression required to promote entry into S-phase.

Initial data demonstrated that quiescent 10-2 or 3-5B cells stimulated with serum, PDGF, or zinc dependent myc overexpression responded by traversing G1 and entering S-phase as expected (Armelin et al., 1984; Rapp et al., 1985; Cavalieri and Goldfarb 1987). Within each cell line, c-myc overproducing cells enter S-phase in response to c-myc induction in a dose dependent manner (figure 4a). Surprisingly, 3-5B cells, which make substantially less c-myc protein than 10-2 cells (figure 3), enter S-phase two to three times more efficiently in response to myc induction than do 10-2 cells .

Classically, PDGF is regarded as an inefficient mitogen for 3T3 cells but an efficient inducer of c accumulation (Muller et al., 1984; Roberts et al., 1985). Hc 25 our hands, PDGF by itself stimulated up to 35% of growth arre . cells to reenter the cell cycle

Figure 3. Western analysis of zinc stimulated c-myc prtotein expression from cells cultured in defined medium. NIH 3T3 cells, 10-2 cells, and 3-5B cells were seeded, brought to growth arrest, annd induced as indicated in materials and methods. The labelling antibody was ^{125}I conjugated anti c-myc monoclonal H60C37. **A.** (-): cells were mock induced. (+): 3T3 cells were induced by adjusting the medium to 5% fetal bovine serum. 3-5B and 10-2 cells were induced by adjusting the medium to 20 μM $ZnSO_4$. All cells were harvested after 4 hours of induction. **B.** 3-5B and 10-2 cells were harvested over a 16 hour timecourse following induction. Cells in the "0" timepoint were harvested 4 hours after mock induction. The other samples were harvested at the times indicated at the tops of the lanes.


A
NIH 3T3
3-5B
10-2
– + – + – +
69 kd
46 kd
1 2 3 4 5 6


B
3-5B cells
Time at harvest (hrs)
0 .5 1 2 4 8 16
69 kd
46 kd
10-2 cells
Time at harvest (hrs)
0 .5 1 2 4 8 16
69 kd
46 kd

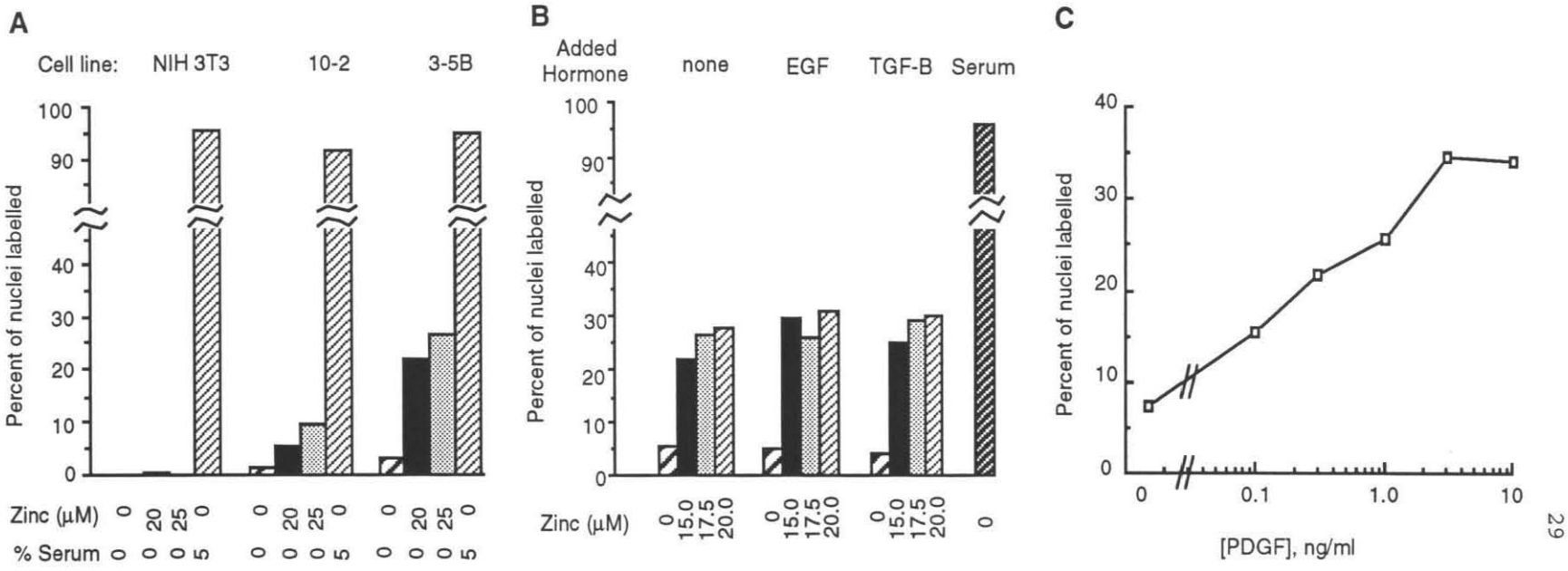


Figure 4. Continuous induction DNA replication assays. In all three experiments, cells were seeded, brought to growth arrest, ^{3}H thymidine labelled, fixed, autoradiographed, and stained as indicated in materials and methods. The labelling window was 14 to 24 hours after zinc or serum treatment. **A.** NIH 3T3 and myc23c expressing 10-2 and 3-5B cells were seeded and brought to growth arrest. Media were adjusted to 0, 20, or 25 μM zinc, or adjusted to 5% FBS. Data are averaged from 3 independent experiments. **B.** 3-5B cells were seeded and brought to growth arrest as above. Media were adjusted to 0, 15, 17.5, or 20 μM with EGF concentrations up to 10 ng/ml or TGF-β concentrations up to 1 ng/ml; or 5% FBS. **C.** NIH 3T3 cells were seeded and brought to growth arrest as above. Cells were mock induced or stimulated with PDGF at the concentrations indicated along the X-axis. Data are averaged from two independent experiments.

(figure 4C). EGF and TGF-B by themselves did not stimulate entry into S-phase. Also, neither EGF nor TGF-B cooperated with either PDGF or zinc stimulated myc expression in this assay (figure 4b). Although we were able to alter growth conditions such that EGF both stimulated cell cycle reentry by itself and cooperated with just sub-threshold concentrations of PDGF (data not shown), we could not define a condition where EGF was inneffective by itself but still able to cooperate with PDGF.

When zinc pulses ranging in length from one hour to continuous exposure were used to induce transient c-myc inductions (figure 5), the number of quiescent cells recruited into S-phase increases with increasing duration of c-myc expression at least through the 14 hour timepoint. In particular, simulating only the burst of high level c-myc expression characteristic of serum stimulated emergence from G_0 is clearly not sufficient to induce these cells to enter S-phase. From prior experiments in which the timing of S-phase was determined, we know that extending the labelling window past 24 hours of induction does not significantly increase the number of cells seen to enter S-phase. Consequently, consistent with the timecourse of its expression following serum stimulation (see figure 7), we conclude that c-myc activity plays an important role through G1, perhaps extending to S-phase.

RNA analysis during cell cycle progression In this cell system, c-myc stimulation is sufficient to recruit up to 30% of growth arrested cells into S-phase, in this respect resembling the effect of PDGF. This raises two substantial molecular level questions. One is whether other regulators, normally appearing with or before c-myc following serum stimulation, are entirely bypassed by c-myc overexpression, or are instead recruited by c-myc. A second is what are the gene regulatory targets of c-myc, either direct or indirect, and which of those are required to induce progression to S-phase.

To begin to address these questions, we examined the expression of genes of established regulatory importance that are subject to induction

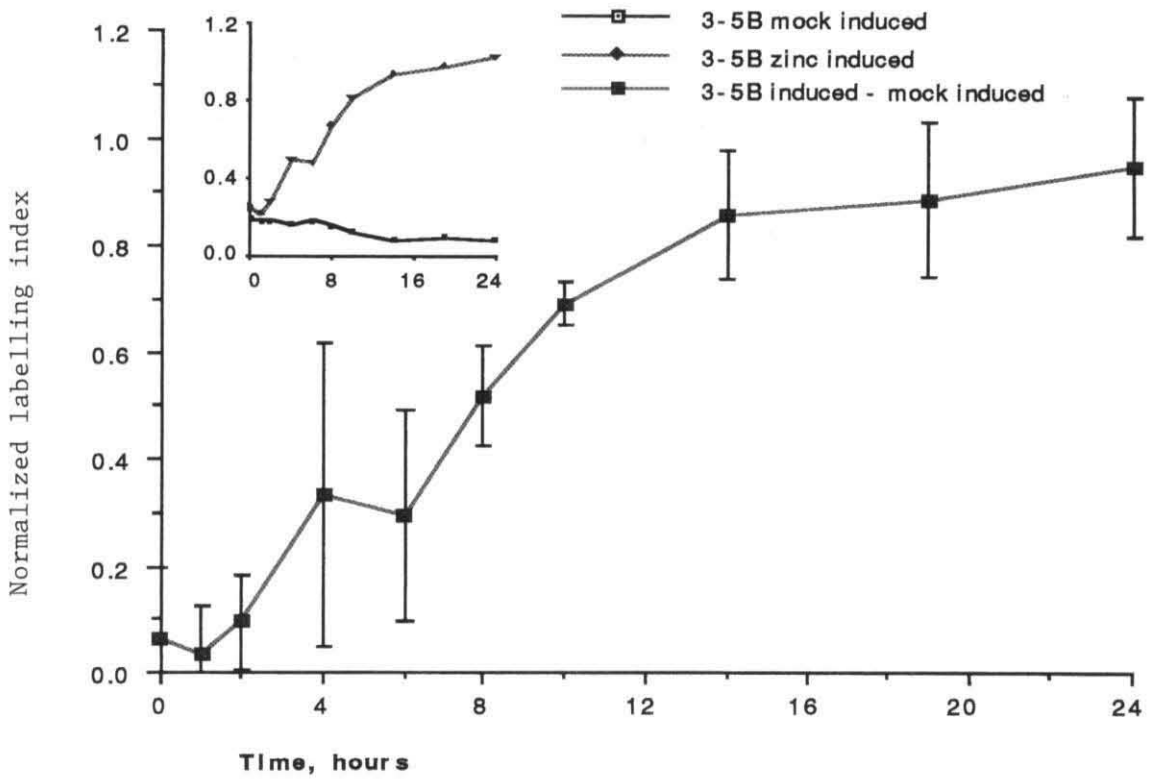


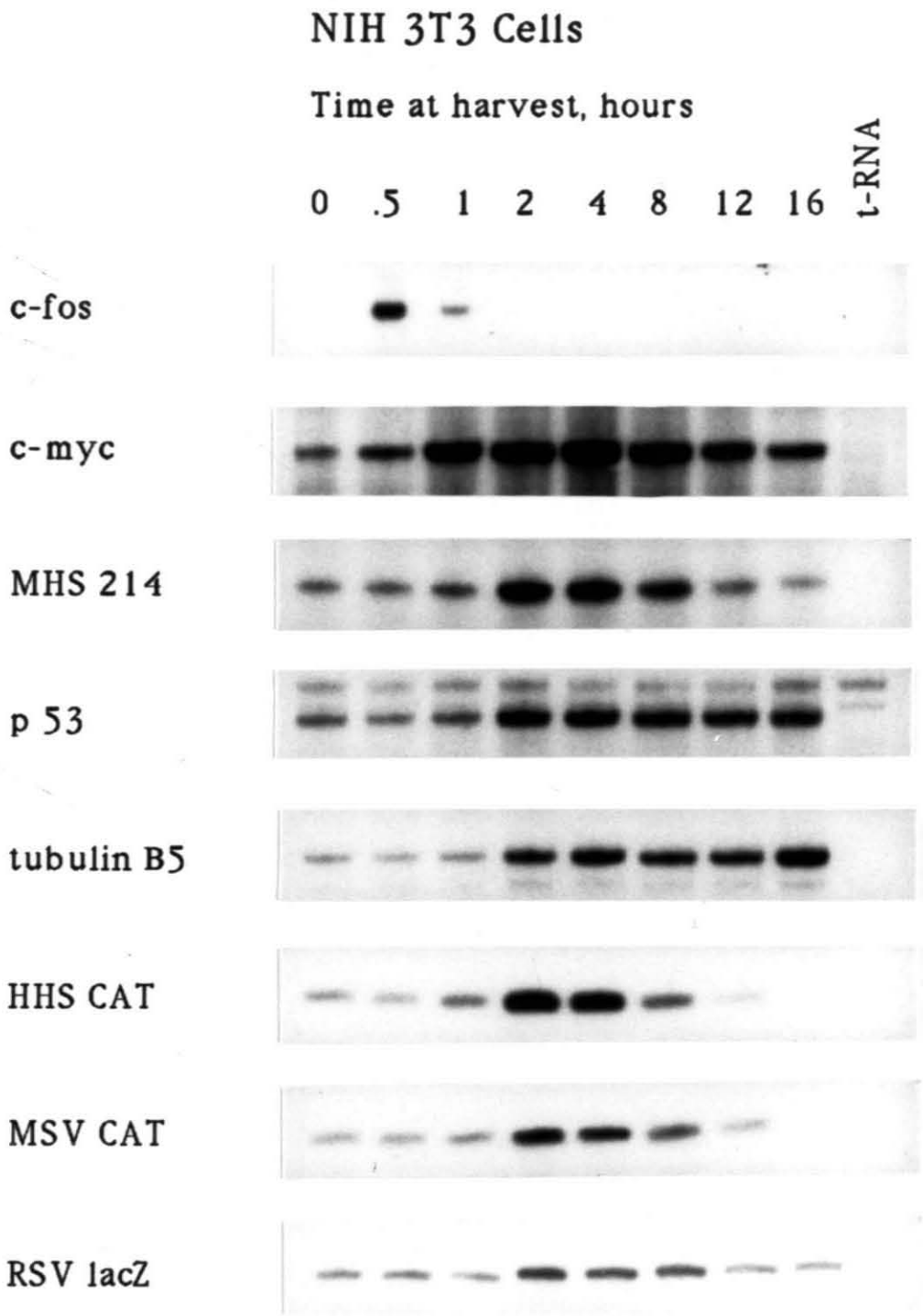
Figure 5. Timecourse of zinc induced DNA replication in 3-5B cells. Cells were seeded and brought to growth arrest as indicated in materials and methods. At t=0, media were adjusted to 20 μM zinc. At the times indicated along the X-axes, zinc containing media were removed, and the cells were washed and refed with defined medium lacking zinc. Cells were ^{3}H thymidine labelled, fixed, autoradioghraphed, stained, and counted as indicated in materials and methods. To combine data from several repetitions of the experiment, each data set was normalized to the average of the induced 19 and 24 hour timepoints within that experiment; each point shown is the average from three independent experiments. Error bars are the sum of one standard deviation for the mock induced data and one standard deviation for the zinc induced data at that timepoint.

following emergence from growth arrest (c-fos, c-myc, p53 tumor antigen), and genes previously reported to be targets of regulation by c-myc (mouse HSP 70, transfected human HSP 70). Since it is not an endogenous mouse gene, Human HSP 70 was examined by transient transfection using MSV CAT and RSV lacZ cotransfected controls. Tubulin β5 again served as an endogenous control.

The expression of our group of test genes was first examined in 3T3 cells that had been growth arrested in defined medium and then stimulated with 5% FBS. An RNA analysis of this experiment is presented in figure 6. In terms of the temporal pattern of their regulation, the transcripts fell into several groups. As expected, c-fos was induced dramatically by the 30 minute timepoint and disappeared within two hours of induction. Also as expected, c-myc was induced by one hour, peaked at four hours, and was still above basal uninduced levels sixteen hours after induction. Endogenous HSP 70, transfected human HSP 70 (HHS CAT), transfected MSV CAT, and transfected RSV lacZ all fell into a third group; the transcripts were induced by two hours after stimulation but had decayed back to basal levels after twelve hours. Finally, the p53 tumor antigen and tubulin β5 transcripts formed yet another group; their transcripts were induced by two hours after stimulation and then stayed at induced levels through the remainder of the timecourse.

Of these, expression of the HSP 70 transcripts were of particular interest because studies from Kingston and colleagues (Kingston et al., 1984; Kaddurah-Daouk et al., 1987) suggested that transiently transfected Drosophila and human HSP 70 promoters might be targets of myc regulation. We first compared the metal response of HSP 70 in parental 3T3 cells and 10-2 cells (figure 7). Basal abundances of the mouse HSP 70 transcripts MHS 213, MHS 214, and MHS 245 (Lowe and Moran 1986) were all on the order of 1-3 transcripts per cell; they were induced sharply by zinc at concentrations above 20 μM. The pattern of expression in parental NIH 3T3 and myc23c expressing 10-2 cell lines was qualitatively similar, the only difference being that the response in

Figure 6. Timecourse nalysis of RNA expression in serum induced NIH 3T3 cells cultured in defined medium. NIH 3T3 cells were seeded, transfected, brought to growth arrest, and induced as indicated in materials and methods. Cells were harvested for RNA preparation over a 16 hour timecourse following induction. Cells in the "0" timepoint were mock-induced and harvested 4 hours later. The other samples were induced by adjusting the media to 5% FBS; cells were harvested at the times indicated at the tops of the lanes. The samples for the 16 hour HHS CAT and 16 hour MSV CAT points were lost. tRNA lane: hybridization containing yeast tRNA but no sample RNA. The analysis was by RNase protection; data on the probes used to detect the indicated messages are presented in materials and methods, table 1.


NIH 3T3 Cells
Time at harvest, hours
0
.5
1
2
4
8
12
16
t-RNA
c-fos
c-myc
MHS 214
p 53
tubulin B5
HHS CAT
MSV CAT
RSV lacZ

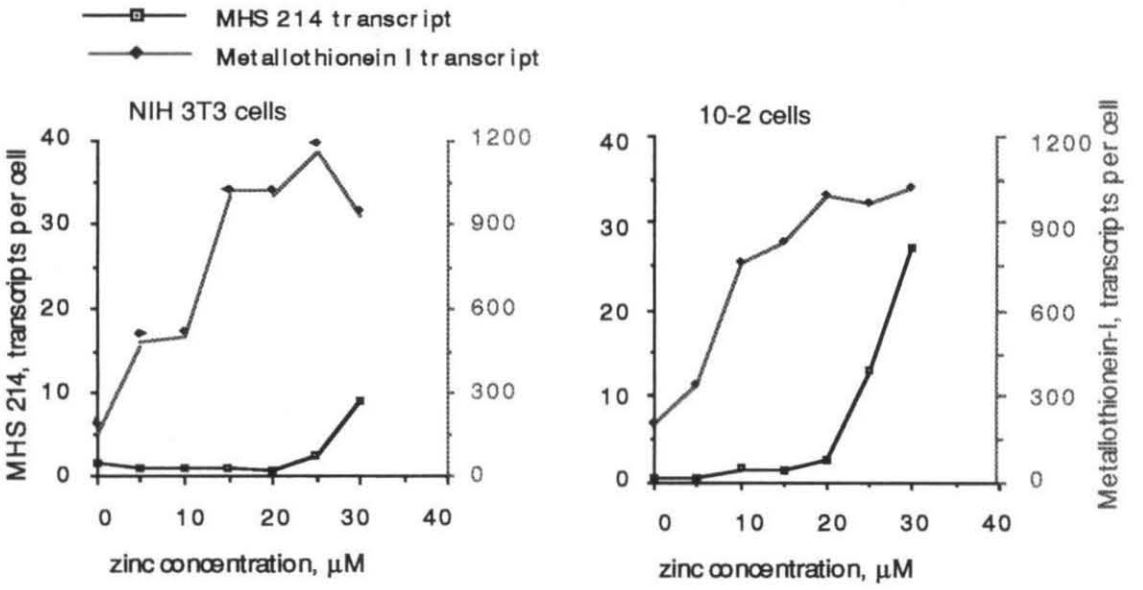


Figure 7. Analysis of HSP 70 and metallothionein-I expression in zinc induced cells cultured in defined medium. Mouse HSP 70 MHS 214 is shown, but MHS 213 and MHS 245 behave similarly. NIH 3T3 cells and myc23c expressing 10-2 cells were seeded and brought to growth arrest as indicated in materials and methods. The media were then adjusted to the zinc concentrations indicated along the X-axes. Cells were harvested for RNA preparation after four hours of induction. Transcript prevalence data were calculated from densitometer scans of autoradiograms of RNase protections. Transcripts per cell was calculated as:

$$\frac{\text{protected cpm } (\mu g \text{ total RNA}^{-1})}{\text{protected probe activity, cpm mol}^{-1}} \times \frac{1\ \mu g \text{ total RNA}}{50{,}000 \text{ cells}} \times 6.02 \times 10^{23} \text{ molecules mole}^{-1}$$

corrected for probe decay over the course of the experiment. CPM represented by each band was backcalculated from an internal standard on the gel. More accurately, this is a measure of transcripts per 20 pg of total RNA; in our experience, a reasonable estimate of the RNA content of growth arrested 3T3 cells.

10-2 cells was shifted to slightly lower zinc concentrations, resulting in larger fold inductions at high zinc concentrations. In light of the amount of c-myc induced in 10-2 cells as the zinc concentration is increased form 0 to 20 μM, the similarity in the pattern of expression of HSP 70 in the two cell lines does not suggest regulation by c-myc. Work by Morimoto and colleagues (Wu et al., 1986) also suggests that HSP 70 transcription can be induced by zinc through the function of a metal response element (MRE) in the promoter; this superficially implies that the two types of transcripts should exhibit similar zinc dose response curves. However, induction of HSP 70 RNA did not parallel that of Mt-I; instead, in both cell lines, HSP 70 accumulation was abruptly induced at zinc concentrations which were near the maximum of the more gradual dose response curve of Mt-I. Although the difference between the zinc dose response curves of the Mt-I transcript and the mouse HSP 70s could be accounted for by very different transcription factor binding affinities and cooperativity at the two types of promoters, a more likely hypothesis is that physiologic stress resulting from high zinc concentrations induces HSP 70 transcription through activation of the heat shock transcription factor.

We proceeded to examine the timecourse of c-myc and HSP 70 expression in 3T3, 10-2, and 3-5B cells (figure 8) at 20 μM zinc, a concentration sufficient to drive about 30% of growth arrested 3-5B cells through G1 into S phase while eliciting little induction of HSP 70. Although the MHS 214 transcript was induced up to 5-fold over the timecourse of the experiment, slightly more than expected on the basis of the experiment in figure 7, its expression did not correlate with that of c-myc; indeed, the largest induction of HSP 70 occurred in the parental 3T3 cells. There was also no correlation between expression of the transfected human HSP 70 transcript and c-myc. Together, these data revealed no regulation of HSP 70 attributable to the effect of c-myc. Since HSP 70 can be induced by zinc alone, we cannot rule out the possibility that a small effect due to c-myc was masked by the larger

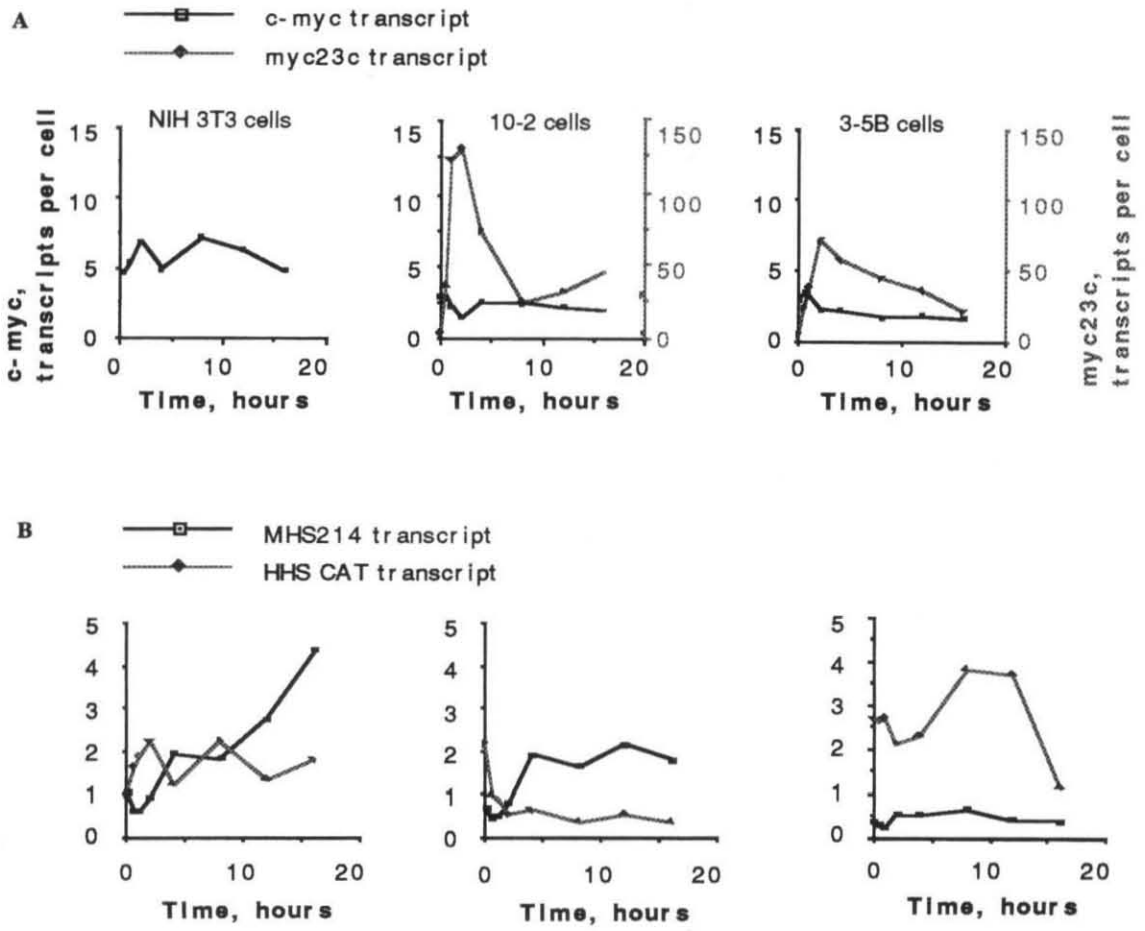


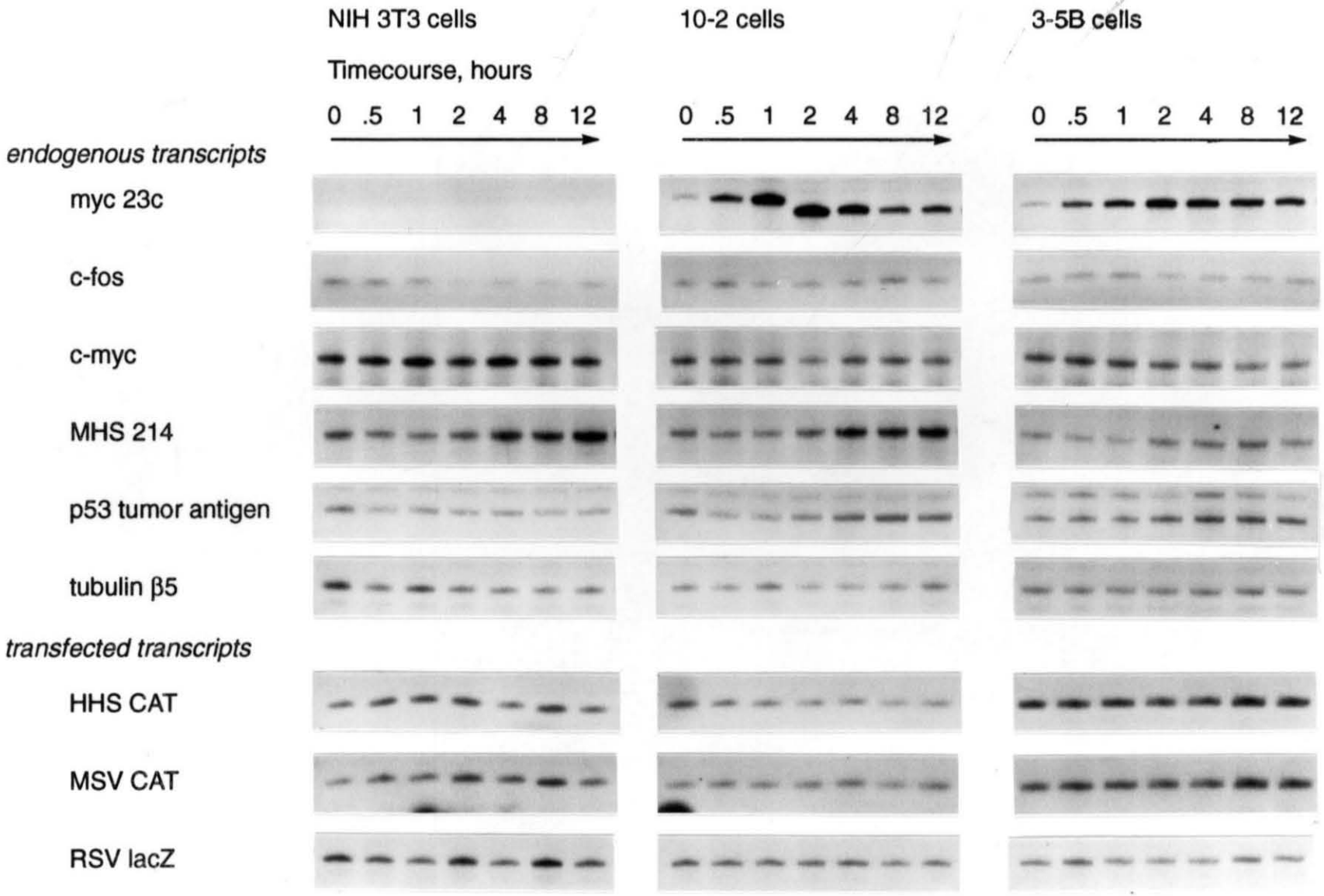
Figure 8. Timecourse analysis of c-myc and HSP 70 transcript expression in zinc induced cells. NIH 3T3, 10-2, and 3-5B cells were seeded, transfected, brought to growth arrest, and induced as indicated in materials and methods. Cells were harvested for RNA preparation over a 16 hour timecourse following induction. Cells in the "0" timepoints were mock induced and harvested 4 hours later; the other samples were harvested at the times indicated along the X-axes.

A. Expression of the endogenous c-myc and transgene myc23c transcripts. Note that the scale for myc 23c is 10-times the scale for c-myc. **B.** Expression of the endogenous HSP 70 MHS 214 and transfected HSP 70 MHS CAT transcripts. Since "transcripts per cell" is rather meaningless for a transiently transfected gene, HHS CAT is plotted on a dimensionless scale. Transcript prevalence data were calculated from densitometer scans of autoradiograms of RNase protections.

effect of zinc; however, we prefer the hypothesis that previous reports of HSP 70 regulation by c-myc were artifactual consequences of cell cycle and stress responses resulting from transient calcium phosphate cotransformations.

Expression of c-fos, p53 tumor antigen, tubulin β5, and transfected MSV CAT transcripts were also measured from the RNA samples used in the experiment of figure 8 (figure 9). In all 3 cell lines, c-fos RNA was present at less than 1 transcript per cell and did not vary significantly with zinc addition. This datum demonstrates that the burst of c-fos RNA synthesis following serum stimulation of growth arrested cells is bypassed upon c-myc stimulated cell cycle reentry. As in figure 2, endogenous c-myc RNA levels were repressed in 10-2 and 3-5B cells relative to 3T3 cells; they were further repressed over the timecourse of the experiment in 3-5B cells. While p53 tumor antigen and tubulin β5 were co-regulated during serum stimulation, they were not co-regulated here. P53 tumor antigen transcript levels remained steady in control 3T3 cells, but they were induced 2- to 3-fold between 4 and 12 hours after zinc addition in 10-2 and 3-5B cells; expression of tubulin β5 RNA was near constant in all three cell lines. Induction of p53 tumor antigen here was not completely unexpected; its timecourse of induction following serum stimulation of growth arrested 3T3 cells is consistent with the possibility that c-myc might play a regulatory role, and p53 tumor antigen is among those proteins which are often overexpressed in cells whose growth characteristics have been altered by c-myc overexpression (Reich and Levine 1984; Shiroki et al., 1986). Expression from the transfected human HSP 70 and MSV CAT were quite similar to each other and did not vary with zinc addition in any consistent manner.

Figure 9. Timecourse analysis of transcript expression from zinc induced cells cultured in defined medium. NIH 3T3, 10-2, and 3-5B cells were seeded, transfected, brought to growth arrest, and induced as indicated in materials and methods. Cells were harvested for RNA preparation over a 16 hour timecourse following induction. Cells in the "0" timepoint were mock induced and harvested 4 hours later. The other samples were induced by adjusting the media to 20 μM $ZnSO_4$; cells were harvested at the times indicated at the tops of the lanes. *: the "0" timepoint of the c-myc NIH 3T3 panel was lost. In other experiments, this point was indistinguishable from adjacent points. tRNA lane: hybridization containing yeast tRNA but no sample RNA. The analysis was by RNase protection.


NIH 3T3 cells
10-2 cells
3-5B cells
Timecourse, hours
0 .5 1 2 4 8 12
0 .5 1 2 4 8 12
0 .5 1 2 4 8 12
endogenous transcripts
myc 23c
c-fos
c-myc
MHS 214
p53 tumor antigen
tubulin β5
transfected transcripts
HHS CAT
MSV CAT
RSV lacZ

Discussion

We have prepared and characterized two cell lines that express a metallothionein promoted c-myc gene. Following treatment with zinc or cadmium, induction of endogenous metallothionein I and stably transfected myc23c RNA occurred as expected. C-myc protein accumulated in the nuclei, and the cells, if they had been growth arrested, reentered the cell cycle, traversed G1, and entered S-phase.

The complexity of the immediate early serum response, the number of known or suspected transcription factors regulated by that response, and the documented changes in both transcriptional and post transcriptional regulation occurring with that response (Lau and Nathans 1987; Almendral et al., 1988), reasonably imply that c-myc acts in concert with other mid G1 regulatory factors. In this light it is not surprising that c-myc overexpression by itself is only modestly effective in stimulating growth arrested cells to progress through G1 to S-phase.[2] The differences between individual cells that do or do not initiate DNA synthesis in response to c-myc, the number of obligatory steps lying between the role(s) of c-myc and initiation of DNA synthesis, and the relative importance of regulatory pathways perhaps parallel or perhaps interconnected with those involving c-myc, are all presently unknown. While our approach can identify some regulatory effects initiated by c-myc alone, those which are co-dependent upon c-myc and other cell cycle regulated proteins are less likely be observed.

First, previous work in constitutively overexpressing cell lines has demonstrated negative autoregulation by c-myc (Leder et al., 1983; Rapp et al., 1985; Mango et al., 1989); that effect is observed here. Both 10-2 and 3-5B cells express less endogenous c-myc transcript by about 2-fold relative to NIH 3T3 cells. 3-5B cells, which show a stronger cell cycle response to c-myc overproduction than 10-2 cells,

---

[2] It remains to be seen whether coexpressed c-myc and max/myn will prove significantly more effective than c-myc by itself at recruiting growth arrested cells back into the cell cycle. However, max/myn is subject to delayed early serum response (Prendergast et al., 1991); unless it turns out to be positively regulated by c-myc, its activity should become limiting in growth arrested cells which have been induced to overproduce c-myc.

further repress their endogenous c-myc transcript following zinc induction. The kinetics of the response, first visible at the 2 hour timepoint when c-myc protein levels also peak, suggest that it might be relatively direct.[3]

Previous work with transiently cotransfected cells suggested that HSP 70s were targets of regulation by c-myc. In our work, no purely myc dependent induction of either endogenous or transfected HSP 70 genes could be observed. Comparison of mouse HSP 70 and transfected human HSP 70 responses to zinc induction in parental 3T3, 10-2 and 3-5B cells at zinc levels sufficient to stimulate progression to S-phase in the myc23c expressing cell lines (figure 8) revealed no effect attributable to c-myc. However, since HSP 70 can clearly be induced by zinc alone, probably through stress related activation of the heat shock transcription factor, we cannot rule out the possibility that a small effect due to c-myc was masked by the larger effect of zinc.

Comparison of p53 tumor antigen and tubulin β5 expression provides an example of differential regulatory responses to serum and c-myc. Both transcripts respond to serum with similar kinetics. P53 tumor antigen does not respond to zinc in 3T3 cells, but it is induced in the 4 to 12 hour time range following induction of c-myc. In contrast, tubulin β5 responds to neither zinc nor c-myc. Thus the similarity of the serum response by p53 and tubulin β5 can be dissected by their different responses to c-myc. This suggests that p53 is, directly or indirectly, a target of regulation by c-myc, while the serum regulated pathway leading to induction of tubulin β5 cannot be directly myc dependent. Although reproducible, induction of the p53 tumor antigen transcript is relatively modest and delayed with respect to the peak of c-myc levels. On the basis of the kinetics and magnitude of its response, it seems most likely that p53 is an indirect target, reflecting events in the 20-30% of cells that enter S-phase in response to c-myc expression.

On the basis of previously published work, we had expected that EGF would augment the cell cycle effect, and presumably the gene

---

[3] Negative autoregulation of c-myc occurs at the level of transcription initiation, but the identity(ies) of the transcription factor(s) involved has not yet been determined (Grignani et al., 1990; Penn et al., 1990).

regulatory effects, of c-myc overexpression. Although we were able to observe an interaction between just sub-threshold concentrations of PDGF and EGF on quiescent 3T3 cells, no cooperative interaction between c-myc and either EGF or TGF-β was found. The difference between our work and previous studies likely arises from differences in the cell lines, culture conditions, or purity of growth factor preparations .

In this work we confirmed expected negative autoregulation of c-myc and provided evidence that the p53 tumor antigen transcript is a target of regulation by c-myc. Working with the hypothesis that the activity of c-myc is responsible for regulating an identifiable subset of the genes induced or repressed in G1 following emergence from growth arrest, our data suggest that screening a large number of cDNA clones subject to mid G1 induction or repression in similar fashion would identify other likely targets of c-myc. While for any individual clone there would be no *a priori* way of knowing how many regulatory steps occur between c-myc overproduction and an observed accumulation or decay of transcript abundance, isolation of a substantial number of such clones should permit definition of one or more pathways involving c-myc. Identification of such target genes will be central in understanding the biological role played by c-myc in regulating the normal cell cycle.

References

Allen-Hoffman, B. L., Crankshaw, C. L., and Mosher, D. F. (1988). Transforming growth factor B increases cell surface binding and assembly of exogenous (plasma) fibronectin by normal human fibroblasts. Molecular and Cellular Biology **8**, 4234-4242.

Almendral, J. M., Sommer, D., MacDonald-Bravo, H., Burckhardt, J., Perera, J., and Bravo, R. (1988). Complexity of the early genetic response to growth factors in mouse fibroblasts. Molecular and Cellular Biology **8**, 2140-2148.

Armelin, H. A. M., Armelin, C. S., Kelly, K., Stewart, T., Leder, P., Cochran, B. H., and Stiles, C. D. (1984). Functional role for c-myc in mitogenic response to platelet-derived growth factor. Nature **310**, 655-660.

Bernard, O., Cory, S., Gerondakis, S., Webb, E., and Adams, J. M. (1983). Sequence of the murine and human cellular myc oncogenes and two modes of myc transcription resulting from chromosome translocation in B lymphoid tumours. European Molecular Biology Organization Journal **2**, 2375-2383.

Bishop, J. M. (1986). Viral Oncogenes. Cell **42**, 23-38.

Blackwood, E. M., and Eisenman, R. N. (1990). Max: A helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with myc. Science **251**, 1211-1217.

Bond, V. C., and Wold, B. (1987). Poly-l-ornithine-mediated transformation of mammalian cells. Molecular and Cellular Biology **7**, 2286-2293.

Caudy, M., Vassin, H., Brand, M., Tuma, R., Jan, L. Y., and Jan, Y. N. (1988). Daughterless, a Drosophila gene essential for both neurogenesis and sex determination, has sequence similarities to myc and the achaete-scute complex. Cell **55**, 1061-1067.

Cavalieri, F., and Goldfarb, M. (1987). Growth-factor-deprived BALB/c 3T3 murine fibroblasts can enter the S phase after induction of c-myc gene expression. Molecular and cellular biology **7**, 3554-3560.

Cavalieri, F., and Goldfarb, M. (1988). N-myc proto-oncogene expression can induce DNA replication in balb/c 3T3 fibroblasts. Oncogene **2**, 289-291.

Cole, M. D. (1986). The myc oncogene: its role in transformation and differentiation. Annual Review of Genetics **20**, 361-384.

Drabent, B., Genthe, A., and Benecke, B. J. (1986). In vitro transcription of a human hsp 70 heat shock gene by extracts prepared from heat-shocked and non-heat-shocked human cells. Nucleic Acids Research **14**, 8933-8949.

Dyson, N. K., Buchkovich, K., Whyte, P., and Harlow, E. (1989). The cellular 107K protein that binds to adenovirus E1A also associates with the large T antigens of SV40 and JC virus. Cell **58**, 249-255.

Ewen, M. E., Ludlow, J. W., Marsilio, E., Decaprio, J. A., Millikan, R. C., Cheng, S. H., Paucha, E., and Livingston, E. M. (1989). An N-terminal transformation-governing sequence of SV40 large T antigen contributes to the binding of both p110Rb and a second cellular protein p120. Cell **58**, 257-267.

Figge, J., Webster, T., Smith, T. F., and Paucha, E. (1988). Prediction of similar transforming regions in simian virus 40 large T, adenovirus E1A, and myc oncoproteins. Journal of Virology **62**, 1814-1818.

Gregor, P. D., Sawadogo, M., and Roeder, R. G. (1990). The adenovirus major late transcription factor USF is a member of the helix-loop-helix group of regulatory proteins and binds DNA as a dimer. Genes and Development **4**, 1730-1740

Grignani, F., Lombardi, L., Inghirami, G., Sternas, L., Cechova, K., and Dalla-Favera, R. (1990). Negative autoregulation of c-myc gene expression is inactivated in transformed cells. European Molecular Biology Organization Journal **9**(12), 3913-3922.

Harland, R., and Weintraub, H. (1985). Translation of mRNA injected into Xenopus oocytes is specifically inhibited by antisense RNA. Journal of Cell Biology **101**, 1094-1099.

Heikkila, R., Schwab, G., Wickstrom, E., Loke, S. L., Pluznik, D. H., Watt, R., and Neckers, L. M. (1987). A c-myc antisense oligodeoxynucleotide inhibits entry into S-phase but not progress from G_0 to G_1. Nature **328**, 445-449.

Hoeffler, W. K., Kovelman, R., and Roeder, R. G. (1988). Activation of transcription factor IIIC by the adenovirus E1A protein. Cell **53**, 907-920.

Hu, Y.-F., Luscher, B., Admon, A., Mermod, N., and Tjian, R. (1990). Transcription factor AP-4 contains multiple dimerization domains that regulate dimer specificity. Genes and Development **4**, 1741-1752.

Jones, T. R., and Cole, M. D. (1987). Rapid cytoplasmic turnover of c-myc mRNA: requirement of the 3' untranslated sequences. Molecular and Cellular Biology **7**, 4513-4521.

Kaddurah-Daouk, R., Greene, J. M., Baldwin, A. S., and Kingston, R. E. (1987). Activation and repression of mammalian gene expression by the c-myc protein. Genes and Development **1**, 347-357.

Kingston, R. E., Baldwin, A. S., and Sharp, P. A. (1984). Regulation of heat shock protein 70 gene expression by c-myc. Nature **312**, 280-282.

Lassar, A. B., Davis, R. L., Wright, W. E., Kadesch, T., Murre, C., Voronova, A., Baltimore, D., and Weintraub, H. (1991). Functional activity of myogenic HLH proteins requires hetero-oligerization with E12/E47-like proteins in vivo. Cell **66**(2), 305-315.

Lau, L. F., and Nathans, D. (1987). Expression of a set of growth-related immediate early genes in Balb/c 3T3 cells: coordinate regulation with c-fos or c-myc. Proceedings of the National Academy of Sciences USA **84**, 1182-1186.

Leder, A., Pattengale, P. K., Kuo, A., Stewart, T. A., and Leder, P. (1986). Consequences of widespread deregulation of the c-myc gene in transgenic mice: multiple neoplasms and normal development. Cell **45**, 485-495.

Leder, P., Battey, J., Lenoir, G., Moulding, C., Murphy, W., Potter, H., Stewart, T., and Taub, R. (1983). Translocations among the antibody genes in human cancer. Science **222**, 765-771.

Lewis, S. A., Lee, M. G. S., and Cowan, N. J. (1985). Five mouse tubulin isotypes and their regulated expression during development. Journal of Cell Biology **101**, 305-321.

Lowe, D. G., and Moran, L. A. (1986). Molecular cloning and analysis of DNA complementary to three mouse M_r = 68,000 heat shock protein mRNAs. The Journal of Biological Chemistry **261**(5), 2102-2112.

Luscher, B., Kuenzel, E. A., Krebs, E. G., and Eisenman, R. N. (1989). Myc oncoproteins are phosphorylated by casein kinase II. European Molecular Biology Organization Journal **8**, 1111-1119.

Mango, S. E., Schuler, G. D., Steele, M. E. R., and Cole, M. D. (1989). Germ line c-myc is not down-regulated by loss or exclusion of activating factors in myc-induced macrophage tumors. Molecular and Cellular Biology **9**(8), 3482-3490.

Maniatis, T., Fritsch, E. F., and Sambrook, J., Ed. (1982). Molecular Cloning. Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory.

Mueller, P. R., Salser, S. J., and Wold, B. (1988). Constitutive and metal-inducible protein:DNA interactions at the mouse metallothionein I promoter examined by in vivo and in vitro fooptrinting. Genes and Development **2**, 412-427.

Muller, R., Bravo, R., Burckhardt, J., and Curran, T. (1984). Induction of c-fos gene and protein by growth factors precedes activation of c-myc. Nature **312**, 716-720.

Murre, C., McCaw, P. S., and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, myoD1, and myc proteins. Cell **56**, 777-783.

Myers, R. M., Rio, D. C., Robbins, A. K., and Tjian, R. (1981). SV40 gene expression is modulated by the cooperative binding of T antigen to DNA. Cell **25**, 373-384.

Payvar, F., DeFranco, D., Firestone, G. L., Edgar, B., Wrange, O., Okret, S., Gustafsson, J. A., and Yamamoto, K. R. (1983). Sequence-specific binding of glucocorticoid receptor to MTV DNA at sites within and upstream of the transcribed region. Cell **35**, 381-392.

Penn, L. J. Z., Brooks, M. W., Laufer, E. M., and Land, H. (1990). Negative autoregulation of c-myc transcription. European Molecular Biology Organization Journal **9**(4), 1113-1121.

Persson, H., Gray, H. E., and Godeau, F. (1985). Growth-dependent synthesis of c-myc-encoded proteins: early stimulation by serum factors in synchronized mouse 3T3 cells. Molecular and Cellular Biology **5**, 2903-2912.

Prendergast, G. C., and Cole, M. D. (1989). Posttranscriptional regulation of cellular gene expression by the c-myc oncogene. Molecular and Cellular Biology **9**, 124-134.

Prendergast, G. C., Lawe, D., and Ziff, E. B. (1991). Association of Myn, the murine homolog of max, with c-myc stimulates methylation-sensitive DNA-binding and ras cotransformation. Cell **65**(3), 395-407.

Rapp, U. R., Cleveland, J. L., Brightman, K., Scott, A., and Ihle, J. N. (1985). Abrogation of IL-3 and IL-2 dependence by recombinant murine retroviruses expressing v-myc oncogenes. Nature **317**, 434-438.

Reich, N., and Levine, A. J. (1984). Growth regulation of a cellular tumour antigen, p53, in nontransformed cells. Nature **308**, 199-201.

Reichel, R., Kovesdi, I., and Nevins, J. R. (1988). Activation of a preexisting cellular factor as a basis for adenovirus E1A-mediated transcription control. Proceedings of the National Academy of Sciences USA **85**, 387-390.

Roberts, A. B., Anzano, M. A., Wakefield, L. M., Roche, N. S., Stern, D. F., and Sporn, M. B. (1985). Type β transforming growth factor: a bifunctional regulator of cellular growth. Proceedings of the National Academy of Sciences USA **82**, 119-123.

Schmid, P., Schultz, W. A., and Hameister, H. (1989). Dynamic expression pattern of the myc protooncogene in midgestation mouse embryos. Science **243**, 226-229.

Schuler, G. D., and Cole, M. D. (1988). GM-CSF and oncogene mRNA stabilities are independently regulated in trans in a mouse monocytic tumor. Cell **55**, 1115-1122.

Shiroki, K., Segawa, K., Koita, Y., and Shibuya, M. (1986). Neoplastic transformation of rat 3Y1 cells by a transcriptionally activated human c-myc gene and stabilization of p53 cellular tumor antigen in the transformed cells. Molecular and Cellular Biology **6**(12), 4379-4386.

Tan, T., Wallis, J., and Levine, A. J. (1986). Identification of the p53 Protein Domain Involved in Formation of the Simian Virus 40 Large T-Antigen-p53 Protein Complex. Journal of Virology **59**, 574-583.

Thompson, C. B., Challoner, P. B., Neiman, P. E., and Groudine, M. (1985). Levels of c-myc oncogene mRNA are invariant throughout the cell cycle. Nature **314**, 363-369.

Van Beveren, C., Straaten, F. v., Curran, T., Muller, R., and Verma, I. M. (1983). Analysis of FBJ-MuSV provirus and c-fos (mouse) gene reveals that viral and cellular fos gene products have different carboxy terminii. Cell **32**, 1241-1255.

Wigler, M., Sweet, R., Sim, G. K., Wood, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S., and Axel, R. (1979). Transformation of mammalian cells with genes from procaryotes and eucaryotes. Cell **16**, 777-785.

Wright, W. E., Sassoon, D. A., and Lin, V. K. (1989). Myogenin, a factor regulating myogenesis, has a domain homologous to myoD. Cell **56**, 607-617.

Wu, B. J., Kingston, R. E., and Morimoto, R. I. (1986). Human HSP70 promoter contains at least two distinct regulatory domains. Proceedings of the National Academy of Sciences USA **83**, 629-633.

Zhan, X., and Goldfarb, M. (1986). Growth factor requirements of oncogene-transformed NIH 3T3 and BALB/c 3T3 cells cultured in defined media. Molecular and Cellular Biology **6**, 3541-3544.

Zimmerman, K. A., Yancopoulos, G. D., Collum, R. G., Smith, R. K., Kohl, N. E., Denis, K. A., Nau, M. A., Witte, O. N., Toran-Allerand, D., Gee, C. E., Minna, J. D., and Alt, F. A. (1986). Differential expression of myc family genes during murine development. Nature **319**, 780-783.

Zinn, K., Dimaio, D., and Maniatis, T. (1983). Identification of two distinct regulatory regions adjacent to the human B-interferon gene. Cell **34**, 223-232.

Chapter 2

Cloning of mid-G_1 serum response genes and identification of a subset regulated by conditional myc expression

Sean V. Tavtigian and Barbara J. Wold

Introduction

When exponentially growing NIH or Balb/c 3T3 cells are deprived of serum growth factors, cell cycle progression halts after mitosis and the cells enter a proliferation arrested state termed G_0. Restoration of mitogens activates signalling cascades that stimulate emergence from G_0 growth arrest and support further cell cycle progression (for review, see Pardee 1989). Exit from G_0 initiates a specialized and temporally protracted G_1 phase.

Progression across the G_0/G_1 boundary is characterized by rapid, robust, protein-synthesis-independent up-regulation of transcription from a set of genes termed the immediate early serum response genes (Lau and Nathans 1987; Almendral et al., 1988; Mohn et al., 1991). Prominent among the immediate early response genes are known or suspected transcription factors of the AP-1, myc, nuclear receptor, and zinc finger families (for review, see Bravo 1990). Their synthesis at the G_0/G_1 transition suggests that they lead a transcriptional regulatory cascade important for progression through G_1. From a cell biological perspective, an important next step is to identify the targets of these regulators and eventually define the individual and collective roles played by these targets in executing progression through G_1 and into S-phase. From a more molecular perspective, further understanding of the immediate early regulators centers on learning how they stimulate or repress expression of their target genes. Knowledge of physiological targets of the immediate early regulators is presently very limited, and one goal of this work was to expand the repertoire of candidate physiological targets. The paucity of targets has been especially acute for myc family regulators, and these are of particular interest to us. Consequently, a second goal of this work was to identify genes affected by conditional myc expression in the context of $G_0 \rightarrow S$ progression.

c-myc contains a prominent region of sequence similarity with other transcription factors, the basic/helix-loop-helix/leucine repeat

(bHLHzip) motif. In other bHLHzip proteins, this motif governs protein:protein oligomerization and DNA binding, and its presence suggests by analogy that myc functions as a component of an oligomeric transcription factor. This prediction was supported when it was shown that the bHLHzip bearing C-terminal fragment of c-myc can, at very high concentrations, form homodimers with DNA binding specificity for the E box sequence CACGTG (Blackwell et al., 1990). In addition, fusion of the c-myc basic region, its candidate DNA binding domain, with the HLH dimerization domain of E12 resulted in a chimeric protein that efficiently formed homodimers with myc-type DNA binding specificity (Prendergast and Ziff 1991). A major recent breakthrough was the cloning of a natural pairing partner of c-myc, max/myn protein (Blackwood and Eisenman 1990; Prendergast et al., 1991). Like c-myc, max/myn is a bHLHzip protein. The two proteins form DNA binding heterodimers efficiently *in vitro* and are associated in vivo (Wenzel et al., 1991; Blackwood et al., 1992). Functional studies of max/myn are at a very early stage, but the presence of max/myn has recently been shown to augment the activity of c-myc in the activated ras cotransformation of secondary rat embryo fibroblasts to the growth transformed phenotype, although max/myn alone cannot (Prendergast et al., 1991).

In an immortalized but non-tumorgenic cell environment such as that of NIH or Balb/c 3T3 cells, the patterns of immediate early expression of c-myc and delayed early expression of max/myn (Prendergast et al., 1991) suggest that the activity of a putative c-myc:{max/myn} heteromeric transcription factor will be stimulated by serum and will peak during the extended G_1 that follows. This leads to the hypothesis that myc, in association with max/myn or some similar partner, is responsible for regulating an identifiable subset of the genes that are induced or repressed following serum stimulated emergence from growth arrest. Moreover, in the absence of serum factors, ectopic conditional expression of myc is sufficient to drive G_0 arrested cells through G_1

and into S-phase (Armelin et al., 1984; Cavalieri and Goldfarb 1987; Eilers et al., 1991), implying that among targets of immediate early regulators those that can respond to myc are functionally sufficient to drive progression. Following these lines of reasoning, we have prepared cDNA libraries from growth arrested and mid-G_1 serum stimulated 3T3 cells, differentially screened those libraries with corresponding cDNA probes, and obtained a panel of cDNA clones subject to either up- or down-regulation with mid-G_1 kinetics. Previously, we had constructed cell lines that conditionally overproduce c-myc in response to treatment with zinc and, following growth arrest, progress through G_1 in response to c-myc overproduction (see chapter 1). In combination with conceptually similar cell lines expressing a c-myc estrogen receptor fusion (Eilers et al., 1989; Eilers et al., 1991), these cell lines here served as sources of RNA with which to screen the panel of serum regulated clones for their response to myc under conditions where myc overproduction drives progression from G_0 through S-phase.

Materials and Methods

Recombinant plasmids used in stable transfections. Plasmids were constructed essentially according to Maniatis (Maniatis et al., 1982). Plasmid myc23c encodes the mouse c-myc proto-oncogene and was used to stably transfect NIH 3T3 cells. The plasmid was assembled in several steps in pT7T3-18 (Bethesda Research Laboratories, Gaithersburg, MD). From 5' to 3', it consists of a 756 bp Taq I - Sac I fragment from the MTV LTR bearing a glucocorticoid regulated enhancer (Payvar et al., 1983), a 218 bp Sac I - Bgl II fragment containing the mouse metallothionien I promoter (Mueller et al., 1988), a 550 bp Xba I - Sac II fragment from the first intron and second exon of the mouse c-myc genomic clone lambda myc 5 (Hood and Barth, unpublished), a 922 bp Sac II - Nsi I fragment containing the remainder of the c-myc coding sequence derived from the c-myc cDNA myc 29 (Kim, Mueller, and Wold, unpublished), and a 302 bp Hgi AI - Hind III fragment containing the poly-A addition signal from the mouse Mt-1 gene (Mueller et al., 1988). The Taq I site at the 5' end of the enhancer fragment was filled-in and ligated to a filled-in Eco RI site within the polylinker, regenerating the Eco RI site. The Bgl II site at the 3' end of the Mt-I promoter was ligated into the Bam HI site of the polylinker, destroying both sites; similarly, the Nsi I - Hgi AI fusion between the 3' end of the myc cDNA fragment and the 5' end of the Mt-I poly-A addition fragment destroyed both sites.

Plasmid pNeo 3, an HSV tk promoted neo construct (Bond and Wold 1987), was used as a G418 selectable marker in the preparation of stably transfected cell lines.

Cell culture. At the outset of this work, NIH 3T3 cells and c-myc transfected derivitives were maintained in Dulbecco's modified Eagle medium (DME) supplemented with 10% fetal bovine serum (Hyclone Laboratories), 100 U of penicillin G per ml, and 7.5 U of streptomycin

per ml. Cells were transfected with a calcium phosphate co-precipitate of plasmid and carrier DNAs by a standard procedure (Wigler et al., 1979). Using pNeo 3 as a selectable marker, NIH 3T3 cells co-transfected with pmyc23c were initially selected for growth in 600 µg/ml G418 (Gibco). After 3 days, the G418 concentration was reduced to 400 µg/ml, and individual clones were maintained in 200 µg/ml G418. In the first two sets of transfections, relatively fast growing clones were chosen for individual analysis, while remaining clones were pooled. Of more than fifteen individual clones checked, none synthesized more RNA from the myc23c transgene than from their endogenous c-myc gene. In contrast, the pooled cell lines expressed considerably more myc RNA. In a third transfection, slow growing clones were deliberately selected. Of 10 clones tested, 4 expressed high levels of the myc23c mRNA. Initial analysis of RNA from these cells using probes sensitive to the 5' and 3' ends of the myc23c transcript showed that the predicted transcript was made. Subsequently, several individual myc23c expressing lines were reisolated from the myc23c expressing pools.

For cell cycle progression work and RNA preparation, NIH 3T3, c-myc transfected derivitives thereof, and Balb/c 3T3 cells were maintained in DME with 15 mM hepes (Irvine Scientific) supplemented with 10% calf serum (Gibco) and 1x glutamine penn strep (Irvine Scientific). For experiments comparing Balb/c 3T3 cells with Balb/c 3T3 derived c-myc estrogen receptor fusion protein expressing MER #6 cells, both were maintained in phenol red free DME supplemented with 5% calf serum (Gibco), 2 mM Hepes free acid, and 1x glutamine penn strep (Irvine Scientific).

Culture in defined medium was essentially as in Zhan and Goldfarb (Zhan and Goldfarb 1986), with the modifications noted below. Tissue culture dishes (15 cm and 24 well, Nunc) were first coated with polylysine (1 mg/ml in H_2O, Sigma) for 2-3 hours at 37°, rinsed with phosphate buffered saline (PBSA) , and coated overnight at room temperature with bovine fibronectin (15 µg/ml in PBSA, Sigma). For cell

cycle progression work and RNA preparation, defined medium (DMI-2) consisted of DME F-12 hepes (Flow Laboratories), 4 μM manganese chloride, 10 μM ethanolamine, 100 nM sodium selenite, 1x glutamine (Irvine Scientific), 5 μg/ml apotransferrin, and 500 μg/ml bovine serum albumin-linoleic acid complex (Sigma). Hormonal supplements were 30 ng/ml insulin and 10 pg/ml TGF-β (RD Systems, Inc). We found that low concentrations (10 pg/ml is $1/10^{th}$ the ED_{50} for stimulation of soft agar colony formation by NRK fibroblasts) of TGF-β, which should facilitate assembly of exogenous fibronectin (Allen-Hoffman et al., 1988), improved cell survival through the timecourse of the experiments. We did not add histidine or hydrocortisone.

For experiments comparing Balb/c 3T3 cells with MER #6 cells, defined medium (DMI-3) consisted of 1:1 phenol red free DME and phenol red free F12 (Gibco), 10 mM Hepes pH 7.2, 4 μM manganese chloride, 10 μM ethanolamine, 100 nM sodium selenite, 1x glutamine / 0.1x penn strep (Irvine Scientific), 5 μg/ml apotransferrin (Sigma), 500 μg/ml bovine serum albumin-linoleic acid complex, 30 ng/ml insulin, and and 10 pg/ml TGF-β.

RNA Preparation Polylysine/fibronectin coated 15 cm plates were seeded at a density of $3x10^6$ cells per plate in 25 ml of DME plus 10% calf serum. Twelve hours later the cells were washed with PBS and fed with 25 ml of DMI. Twelve hours later the cells were refed with 25 ml of DMI. Twenty four hours later the cells were serum induced with the addition of FBS to 10% final concentration, zinc induced with the addition of various concentrations of zinc sulfate in DME phosphate free hepes (Irvine), or mock induced with the addition of an appropriate volume of DME phosphate free hepes.

When RNA was intended for library or probe preparation, cells were harvested in 5M guanidinium thiocyanate, 50 mM Tris pH 7.5, 12.5 mM EDTA, 2% N-lauryl sarcosine, and 1% 2-mercaptoethanol. Lysates were adjusted to 1.5 μg/ml ethidium bromide and an approximate density of 1.6

g/ml with CsTFA. Lysates were spun at 41,500 rpm for 30 hours in an SW55 rotor, the RNA band was withdrawn, ethanol precipitated twice, and redissolved in TE (Mirkes 1985).

When RNA was intended for RNase protections, cells were harvested in 4M guanidinium thiocyanate, 25 mM sodium citrate pH 7.0, 0.5% N-lauryl sarcosine, and 100 mM 2-mercaptoethanol. 0.1 volumes of 2M sodium acetate pH 4.0, 1 volume of phenol, and 0.2 volumes of chloroform were added sequentially with vigorous mixing following each addition. The preparations were chilled in an ice bath for 15 minutes and then centrifuged at 10,000 rpm for 10 minutes at 4° in an SS34 rotor. The aqueous phase was reextracted with 1 volume of 1:2 phenol:chloroform, precipitated twice, and redissolved in TE (Chomczynski and Sacchi 1987).

cDNA Clone Selection and Management. Cloning: (1) Poly A plus RNAs were prepared from growth arrested and 8 hour serum stimulated Balb/c 3T3 cells. (2) First strand cDNAs were eventually used as probes. Hemimethylated double stranded cDNAs were directionally cloned into the vector λ EXLX+ (Palazzolo et al., 1990). (3) Biotinylated cRNA prepared by transcribing phage library DNAs en masse was hybridized to the first strand RNA; biotin cRNA from the growth arrested library was hybridized to serum stimulated cDNA and biotin cRNA from the serum stimulated library was hybridized to growth arrested cDNA. Heteroduplexes and unhybridized cRNA were removed with the addition of streptavidin followed by phenol extraction. (4) The resultant subtracted cDNAs were ^{32}P labelled by priming with random nonamers and extending with T7 DNA polymerase. A "backprobe" was generated by similarly labelling a limited collection of serum regulated clones plus clones selected in preceeding rounds of library screening [see step 6]. (5) λ libraries were plated at a density of 1000 plaques per 15 cm plate. Three lifts were made from each plate. One lift was probed with the growth arrested subtracted probe, a second lift was probed with the serum stimulated subtracted probe, and a third lift was probed with the

backprobe. Plaques differentially hybridized to by the first and second probes and not labelled by the backprobe were picked for further analysis. (6) cDNA inserts were PCR amplified from primary plaques using primers based on flanking vector sequences. The 3' end primer was ^{32}P labelled, resulting in 3' end labelled cDNAs. After gel purification, fragments were subjected to partial restriction digest mapping from their 3' ends with Hae III, Hin fI, and Sau 3A. Restriction maps were compared by computer analysis (appendix C). Where a mRNA species was represented by more than one clone, only the longest corresponding cDNA was used for further analysis. cDNAs from this level of analysis were combined for use in the backprobe [see step 4]. The mid-G1 library was screened through 6 rounds of 6 plates each, while the growth arrested library was screened for two rounds of 6 plates.

Characterization: (1) cDNA bearing plasmids were excised from primary λ clones by a cre-lox recombination system (Palazzolo et al., 1990). Where more than one cDNA was obtained from the primary plaque, they were all maintaind for further analysis. (2) cDNAs were dideoxy sequenced from their 5' ends. (3) Resultant sequences were used to generate restriction maps as well as for seqence homology searches of GENBANK and EMBL using the program BLAST (Altschul et al., 1990). (4) A 5' end fragment from each cDNA was subcloned into a duel polymerase promoter plasmid to generate a riboprobe template. (5) The timecourse of expression of each RNA in serum stimulated NIH 3T3 cells was checked by ribonuclease protection. (6) Expression of those clones subject to serum regulation with mid-SR kinetics was further characterized in the conditional myc expressing cell lines 3-5B (see chapter 1) and MER #6 (Eilers et al., 1989; Eilers et al., 1991).

PCR cloning. Probes for the transcripts G1 cyclin D2, prothymosin α, and max/myn were cloned by amplification from first strand cDNA. PCR primers were:

Cy D2 5': GCG GAA TTC TGG ATG CT[C/G] GAG GT[A/C/G/T] TG[C/T] GA

Cy D2 3': CGC GGA TCC IGC [A/C/G/T]A[A/C/G/T] GAA GGT CTG IGC ITG [C/T]TT

prothymosin α 5': CGC GAA TTC ATG TCA GAC GCG GC[A/C/G/T] GT[A/C/G/T] GA

prothymosin α 3': CGC GGA TCC CTG CTT CTT GGT [C/T]TC [A/C/G/T]AC [G/A]TC

max/myn 5': GCA GAT CTG CCA CCA TGA GCG ATA ACG ATG ACA T[C/T]G A[G/A]G T

max/myn 3': GCG TCG ACT CAG CTG GCC TCC ATC CGG A[G/A][C/T] TT[C/T] TT

Prothymosin α and G1 cyclin D2 fragments were cloned as Eco RI - Bam HI fragments into duel promoter vectors to serve as riboprobe templates. A Bgl II (PCR generated site) - Sma I fragment from the 5' end of max/myn was subcloned for the same purpose.

Other riboprobe templates. Several other duel promoter plasmids were prepared for use as riboprobe templates in the work. pmyc1-HA contains a 122 bp Hind III - Alu I fragment from the first exon of c-myc. The Hind III site is located between P1 and P2 (Bernard et al., 1983) within the c-myc promoter. pmyc2-XR contains a 192 bp Xba I - Eco RV fragment starting near the 3' end of the first intron of c-myc and extending into the second exon. pMHS-SX contains a 646 bp Sma I - Xba I internal fragment from the mouse HSP 70 cDNA MHS 214 (Lowe and Moran 1986). pfos-PP2 contains a 108 bp Pvu II - Pst I fragment spanning the 5' end of the c-fos transcript (Van Beveren et al., 1983). pp53-XA contains a 1005 bp Xho I - Asp 718 fragment from the p53 tumor antigen gene (Tan et al., 1986). pβ5 contains an approximately 135 bp 3' end fragment from a mouse tubulin β5 cDNA (Lewis et al., 1985) from which the poly-A tract has been trimmed with exonuclease Bal 31.

RNA Measurement. RNA probes were synthesized using T7, T3, or SP6 RNA polymerase essentially as recommended by the enzyme manufacturer. Full length probes were excised from 5% acrylamide gels crosslinked with BAC (N,N'-bis-acrylylcystamine, Bio-Rad) and the acrylamide was dissolved with 200 mM 2-mercaptoethanol in TE. RNase protections were performed by a modification of the procedure described by Zinn et al. (Zinn et al., 1983). One or two µg of total cellular RNA plus tRNA to make a total of 25 µg were coprecipitated with probe. Samples were redissolved in 30 µl of RNA hybridization buffer,(40 mM Pipes pH 6.7, 400 mM NaCl, 1 mM EDTA, 80% formamide) denatured at 80° for 10 minutes, and hybridized for >12 hours. Probe concentrations and hybridization temperatures were optimized so that hybridizations proceeded to at least 90% of completion in 12 hours. Hybridizations were digested with the addition of 300 µl of 1500 Units/ml RNase T1 (BRL) in RNase digestion buffer (10 mM Tris pH 7.5, 300 mM NaCl, 5mM EDTA) for 60 minutes at 30°, followed by 200 µl of 250 µg/ml proteinase K, 1% SDS for 15 minutes at 30°. Handling of samples containing dissolved BAC acrylamide was facilitated by co-precipitating RNA samples and probe out of 100 mM 2-mercaptoethanol and using well silanized tubes. Samples were prepared for electrophoresis, fractionated on 5% denaturing acrylamide gels, dried, and exposed to X-ray film. Count per minute data were also obtained from dried gels on a phosphorimager (Molecular Dynamics) and converted to transcript per cell data using internal CPM standards.

DNA Replication Assays. For experiments with NIH 3T3 and 3-5B cells, Polylysine/fibronectin coated 24 well plates were seeded at a density of $2x10^4$ cells per well in 0.5 ml of DME plus 10% serum. Twelve hours later the cells were washed with PBS and fed with 0.75 ml of DMI. Twelve hours later the cells were refed with 0.75 ml of DMI. Twenty four hours later the cells were zinc induced with the addition of various concentrations of zinc sulfate in DME phosphate free hepes (Irvine), mock induced with the addition of an appropriate volume of DME

phosphate free hepes, or serum induced with the addition of FBS to 5% final concentration. For experiments with Balb/c 3T3 and MER #6 cells, fibronectin coated 24 well plates (Collaborative research) were seeded at a density of 3x104 cells per well in 0.5 ml of phenol red free DME (Gibco) that had been adjusted to 4mM hepes with the addition of dry hepes free acid, 5% calf serum, 1x glutamine penn strep (Irvine Scientific), and 1x GMS-X medium supplement (Gibco). 24 hours later the cells were refed with 0.75 ml of the same medium. 72 hours later the cells were β-estradiol induced with the addition of various concentrations of β-estradiol in phenol red free DME, mock induced with the addition of an appropriate volume of phenol red free DME, or serum induced with the addition of FBS to 15% final concentration. In both procedures, entry into S-phase was monitored by adding 37.5 µl of a labelling mixture containing 5-bromodeoxyuridine and 5-fluorouracil (Amersham) 14 hours after stimulation. Experiments were terminated 10 hours after labelling mix addition by washing the cells withe PBSA and then fixing them in 5% acetic acid in ethanol. Immunoperoxidase staining employed the Amersham cell proliferation kit. Wells were washed 3 x 3 minutes with PBS, incubated with 8-fold diluted anti bromodU mAB / nuclease mix for 4 hours at room temperature, washed 3 x 3 minutes with PBS, incubated with 4-fold diluted peroxidase labelled anti mouse Ab for 2 hours at room temperature, washed 3 x 3 minutes with PBS, stained with nickel/cobalt enhanced DAB, and counterstained with fast green FCF.

Western Blots. Polylysine/fibronectin coated 6 cm plates were seeded at a density of 2.5-4.0 x 10^5 cells per plate in DME plus 10% calf serum. Twelve hours later, plates cells were washed with PBSA and fed with 3 ml of DMI. Twelve hours later cells were refed with 3 ml of DMI. Twenty four hours later cells were induced with 300 µl of 3:1 DME:F12, 150 µl of fetal bovine serum (Hazleton Biologics), or 300 µl of 220 µM $ZnSO_4$ in 3:1 DME:F12. Cells were harvested over a 16 hour

timecourse as indicated in the figure legend. Cells were lysed in 200 µl of loading buffer (80 mM Tris pH 6.8, 100 mM dithiothreitol, 2% SDS, 10% glycerol). Cell lysates were fractionated on acrylamide SDS gels and electroblotted to Nitroplus 2000 (Micron Separations). Filters were blocked for 2 hours at room temperature in 1X TBS, 0.5% Tween 20, 2% BSA. Antibody incubations were in the same solution with the addition of 2 µg/ml anti-myc monoclonal antibody H60C37 ^{125}I labelled to a specific activity of approximately 100 Ci/mMole. Antibody incubations were for 12 hours, followed by five 20 minute washes in blocking solution.

Results

Cloning Strategy. Our goal in this work was to identify genes that are physiological targets of transcriptional regulators induced during the immediate early serum response, especially c-myc. cDNA libraries were prepared from both serum starved and 8-hour serum stimulated Balb/c 3T3 cells. These libraries were screened with probes enriched for sequences either induced or repressed in response to mitogen stimulation. Figure 1 presents an outline of the cloning strategy; detailed considerations are presented in Materials and Methods. The design was influenced by several considerations. First, we focussed the selection scheme on the kinetics of progression through G_1 following emergence from G_0. In this experimental regime, the 8-hour timepoint is reproducibly within G_1, close to the restriction point (R-point), beyond which entry into S-phase becomes insensitive to inhibition by serum removal (Pardee 1989). Thus, execution genes whose new transcription and translation are needed for progression from G_1 to S should be well-represented at the 8-hour point (entry into S under these conditions is at ~14 hours). The 8-hour point was also expected to allow for the decay of prominent G_0 transcripts derived from genes that are down-regulated in response to serum. We will refer to transcripts differentially regulated at 8 hours as mid-serum response, or mid-SR, to distinguish them from a potentially distinct class regulated in mid-G_1 of exponentially growing cells. A second strategic consideration was that we wanted to obtain clones corresponding to mRNAs of all prevalence classes. This was achieved by using subtracted hybridization probes enriched in differentially expressed sequences, which effectively increased the screening sensitivity for lower abundance cDNAs (Sargent 1987; Palazzolo et al., 1990). Third, we anticipated that some genes whose expression is initiated early, qualifying them as members of the immediate early kinetic class, will continue to be expressed through G_1, and that this later phase of

Figure 1

A. c DNA library preparation

1. prepare ds cDNA from 8 hour serum stimulated cells; clone into λ ExLx
   → $15x10^6$ independent clones

2. prepare ds cDNA from G_0 arrested cells; clone into λ ExLx
   → $7x10^6$ independent clones

B. Probe preparation

1. 1^0 probe: ss cDNA derived from serum stimulated cells

2. preparation of subtracting cRNA
   a. prepare λ DNA from arrested library
   b. transcribe with T7 RNA pol and biotinylated NTPs
   → biotinylated sense cRNA

3. preparation of subtracted cDNA
   a. hybridize 1° ss cDNA with biotinylated cRNA
   b. remove heteroduplexes and cRNA
   → serum stimulated subtracted cDNA probe - the unreacted ss cDNA is enriched for sequences differentially up-regulated in seum stimulated cells

C. Library screening

1. plate at low density: ~1000 plaques per plate

2. make 3 lifts per plate:
   a. screen with the serum stimulated subtracted probe
   b. screen with the arrested subtracted probe
   c. screen with the "backprobe" (see below)

3. select clones that:
   a. score positive with the serum stimulated probe
   b. score negative with the arrested probe
   c. score negative with the backprobe
   → low redundancy serum stimulated cDNA clones

D. Preparation of backprobe

1. PCR amplify cDNA inserts from selected clones
2. gel purify and pool DNAs
   → backprobe

expression may also identify them as mid-SR class genes. c-Myc itself, for example, displays such a pattern of expression (Muller et al., 1984; Lau and Nathans 1987). While there are several different possibilities for the regulatory basis of such expression, we wanted to detect such genes in the initial screen. For this reason, we did not initially eliminate all immediate early class genes, but later identified them as a kinetic subset.

An important practical goal in cloning exercises of this sort is to reduce labor by eliminating redundant clones as early in the process as possible. At the level of initial plaque screens, we reduced the repeated selection of cDNAs derived from a few prevalent and differentially expressed mRNAs. As indicated in figure 1, this was done by proceeding through several successive rounds of library screening, making triplicate lifts from each plate. The differential probes were hybridized to the first and second lifts, whereas a "backprobe", generated by pooling cDNAs already obtained in earlier rounds of screening, was hybridized to the third lift; only plaques yielding differential signals on the first and second lifts and scoring negative on the third lift were selected. To further reduce the number of redundant clones maintained after a given round of screening, restriction maps were generated from the 3' ends of the newly selected clones; a computerized comparison of the new clone restriction maps with the existing clone maps (appendix C) was used to identify and omit redundant clones. Plasmids bearing potentially unique cDNAs were automatically subcloned (Palazzolo et al., 1990); plasmid miniprep DNAs were sequenced from their 5' ends, yielding 250 to 325 bp of sequence from each clone. Sequences were searched against the Genbank and EMBL sequence libraries using the program BLAST (Altschul et al., 1990); remaining redundant cDNAs identified during the database searches, usually either short or internally initiated copies of highly expressed mRNAs, were eliminated at this stage.

The growth arrested and serum stimulated cDNA libraries contained $7x10^6$ and $1.5x10^7$ clones, respectively. These cells contain in the range of 3-10 x 10^5 mRNA molecules per cell (Jendrisak et al., 1987), so the libraries are large enough to ensure that even the rarest mRNA species are represented, except for possible exceptional sequences that may have been peculiarly poor substrates for reverse transcription. Although the libraries are comprehensive, the screening reported here was not extensive enough to be saturating and was more complete for the serum induced case than for the repressed case. Thus the serum stimulated library was screened through six rounds of six plates each (at ~1000 plaques per plate) while the growth arrested library was screened through two rounds of six plates each. From 100 differentially hybridized plaques approximately 40 cDNAs were found to be non-redundant by the criteria of 3' end restriction mapping and limited 5' end sequence analysis. These were subjected to further analyses as described below.

Expression in growth arrested and serum stimulated cells. The 40 candidate 8-hour response clones were next tested for the kinetics of their response to serum. At this point members of the immediate early kinetic class were identified and separated from those clones whose RNAs were most strongly differentially regulated at the 8-hour point, mid-G_1 relative to G_0. For operational purposes, we defined mid-SR regulated clones as those either (1) expressed at higher levels 8 hours following serum stimulation than 1 hour following, or (2) repressed during G_1 following serum stimulation of quiescent cells. To confirm magnitude and timing of regulation, ribonuclease protection assays were used. The sources of RNAs used for screening were NIH 3T3 cells that had been brought to growth arrest in defined medium, then mock or serum induced for 1 hour or 8 hours, and harvested for RNA preparation. This screening used RNA from NIH 3T3 cells with the intention of filtering out clones whose response was particular to the Balb/c 3T3 line used to

construct and screen the original libraries. Antisense RNA probes generated from the 5' end of each cDNA were hybridized to total RNA, digested with RNase T1, and fractionated on denaturing polyacrylamide gels. Quantitative data from each set of ribonuclease protections were obtained from dried gels on a phosphorimager (Molecular Dynamics). Using internal standards on each gel, transcript prevalence was calculated (Figure 2). Of the 40 unique clones obtained from initial screens, 18 were classified as mid-SR genes; 15 of these were induced and 3 were repressed by serum treatment. The remaining clones were either subject to stimulation by serum with predominantly early G_1 kinetics or were not reproducibly regulated. The clones expressed most highly at one hour have not yet been subscreened to differentiate those that are regulated independently of protein synthesis (the criterion that distinguishes the immediate early class) from those that are dependent on immediate early products. Nomenclature: serum stimulated clones were initially named I-8-x where I-8 indicates serum induced at 8 hours and x is the number assigned to the primary plaque upon its selection. Repressed clones were named R-8-x where R-8 indicates serum repressed at 8 hours. Where an original phage plug later proved to contain more than one cDNA clone, the clone number is followed by a letter, i.e., I-8-10A and I-8-10B.

Expression of clones subject to mid-SR regulation is presented in figure 2. For comparison with other work, panel A presents the classic immediate early transcripts c-fos and c-jun (Muller et al., 1984; Lau and Nathans 1987; Almendral et al., 1988), as well as the additional recently cloned genes of potential interest max/myn, Id, and G_1 cyclin D2 (Benezra et al., 1990; Blackwood and Eisenman 1990; Lew et al., 1991; Matsushime et al., 1991; Prendergast et al., 1991). The serum induced I-8 clones (panel B) and serum repressed R-8 clones (panel C) correspond to mRNAs from all prevalence classes, as anticipated from the cloning strategy. For example, I-8-16 [cytoplasmic (A-X) actin], R-8-1 [alpha-1 (3) collagen], and R-8-9 [translation elongation factor 1] were all

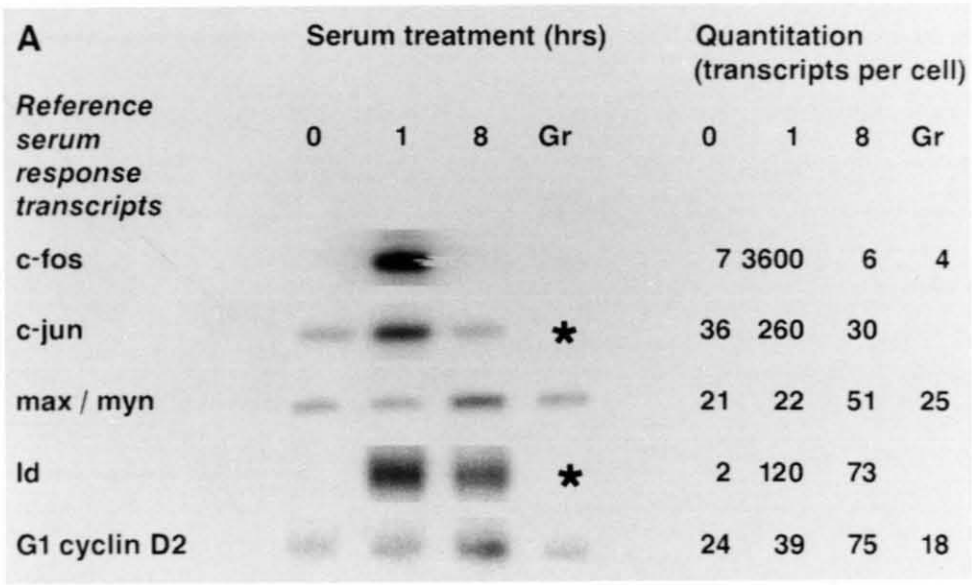
Figure 2. Expression of cloned transcripts regulated in NIH 3T3 cells with mid-SR kinetics.NIH 3T3 cells were seeded and brought to growth arrest in defined medium as indicated in materials and methods. Arrested cells were mock or serum stimulated. RNAs generated from mock, serum stimulated, and exponentially growing cells were analyzed by ribonuclease protection. Quantitative data from each set of ribonuclease protections were obtained from the dried gels on a phosphorimager (Molecular Dynamics). Using internal standards on each gel, transcript per cell was calculated as:

$$\frac{\text{protected cpm } (\mu g \text{ total RNA}^{-1})}{\text{protected probe activity, cpm mol}^{-1}} \times \frac{1\ \mu g \text{ total RNA}}{50{,}000 \text{ cells}} \times 6.02 \times 10^{23} \text{ molecules mole}^{-1}$$

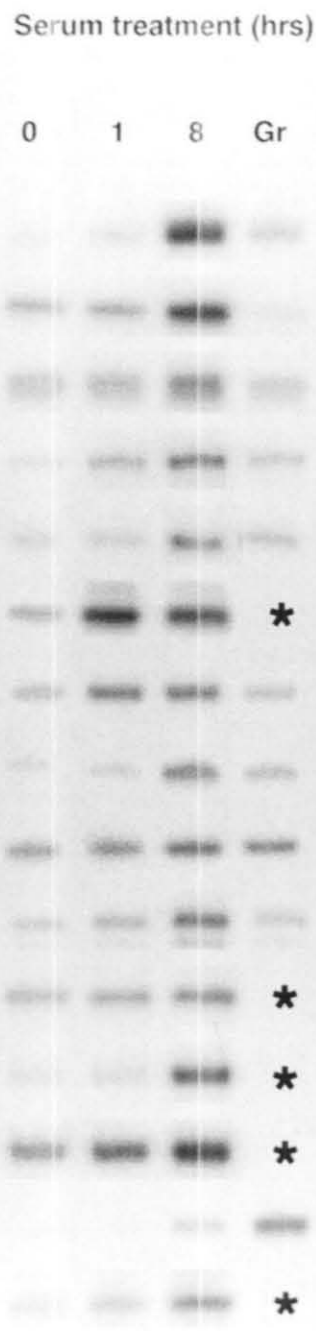
corrected for probe decay over the course of the experiment.

More accurately, this is a measure of transcripts per 20 pg of total RNA; in our experience, a reasonable estimate of the RNA content of growth arrested 3T3 cells.

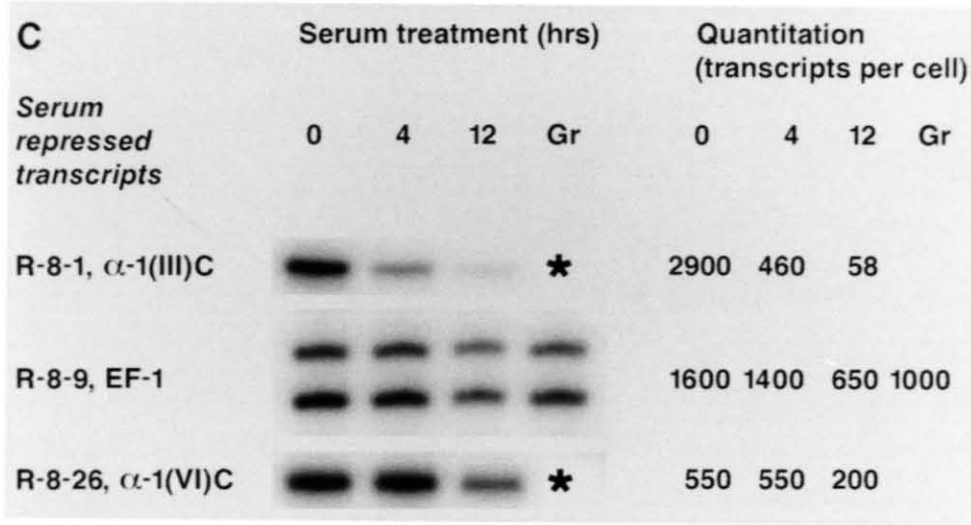
A. Reference serum response transcripts. **B.** Serum induced transcripts. **C.** Serum repressed transcripts. Timecourse: 0, mock induced; 1, 4, 8, 12, serum treated for 1, 4, 8, or 12 hours; Gr, exponentially growing cells. Abbreviations: α-1(III)C, α-1 type III collagen; EF-1, translation elongation factor 1; α-1(VI)C, α-1 type VI collagen; MGSA, melanoma growth stimulatory activity; T1p, T1 protein; LDH, lactate dehydrogenase; LTT, liver thiol transferase; G3PDH, glyceraldehyde-3-phosphate dehydrogenase; ODC, ornithine decarboxylase; HSA, heat stable antigen.


A
Serum treatment (hrs)
Quantitation
(transcripts per cell)
Reference serum response transcripts
0
1
8
Gr
0
1
8
Gr
c-fos
7
3600
6
4
c-jun
*
36
260
30
max / myn
21
22
51
25
Id
*
2
120
73
G1 cyclin D2
24
39
75
18

B



Serum induced transcripts	Quantitation (transcripts per cell) 0	1	8	Gr
I-8-1, tenascin	7	29	880	62
I-8-6, MGSA	47	63	400	11
I-8-9	41	48	143	31
I-8-10A, T1p	10	68	250	39
I-8-10B, LDH	46	65	405	99
I-8-16, actin	1400	6700	7100	*
I-8-23	5	15	19	4
I-8-29	11	13	57	22
I-8-30B	19	24	38	22
I-8-32, LTT	27	45	190	32
I-8-33, G3PDH	440	570	1200	*
I-8-34	11	17	320	*
I-8-36, ODC	32	40	84	*
I-8-41, HSA	22	24	100	240
I-8-47	11	220	790	*


C
Serum treatment (hrs)
Quantitation (transcripts per cell)
Serum repressed transcripts
0
4
12
Gr
0
4
12
Gr
R-8-1, α-1(III)C
*
2900
460
58
R-8-9, EF-1
1600
1400
650
1000
R-8-26, α-1(VI)C
*
550
550
200

present at more than 1000 transcripts per cell. At the other extreme, I-8-23, I-8-29, and I-8-30B were present at 50 or fewer transcripts per cell at the 8 hour serum induction timepoint. The degree of regulation varied from a reproducible but modest 2-fold induction of I-8-30B to more than 100-fold stimulation of I-8-1 [tenascin] and 50-fold repression of R-8-1 [alpha-1 (3) collagen].

The regulatory capacity of conditional myc expression. The burst of transcription factor expression in the immediate early serum response implies a transcriptional cascade acting through the G_1 that follows emergence from G_0. The panel of mid-SR genes identified here are possible direct or indirect targets of the immediate early regulators. We next focused on determining which, if any, could be regulated by c-myc. To screen for myc responsiveness, we used an NIH 3T3 cell line (3-5B) that expresses c-myc driven by the metallothionein-I promoter (see chapter 1). This cell line is subject to growth arrest in defined medium, as is the parental NIH 3T3 line. As in other conceptually similar cell lines with inducible myc activity, c-myc expression under arrest conditions is sufficient to stimulate a subpopulation of the growth arrested cells (50-60%) to reenter the cell cycle, traverse G_1, and enter S-phase (Figure 3). During the course of such an experiment, the transgene c-myc (myc23c) transcript is induced in the 3-5B cell line by about 20-fold (figure 5) and c-myc protein peaks at levels exceeding those present in a serum stimulated NIH 3T3 cells by 5- to 10-fold (figure 4).

The degree of cell cycle progression due to conditional myc expression in this system is very similar to that reported in conceptually similar systems. Thus following growth arrest in defined medium, 40-50% of Balb/c 3T3 cells expressing c-myc under control of a glucocorticoid regulated promoter reenter the cell cycle following transcriptional induction of myc expression (Cavalieri and Goldfarb 1987). And following growth arrest by contact inhibition, 50-60% of

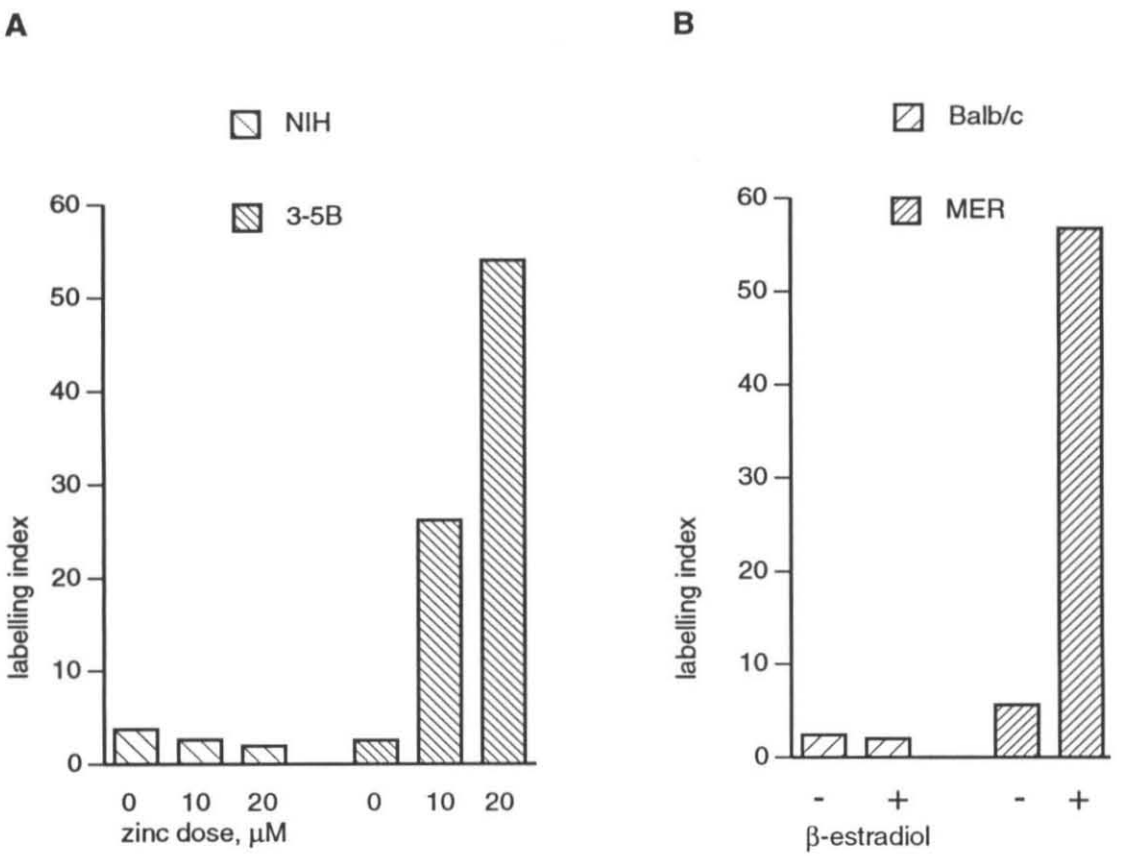
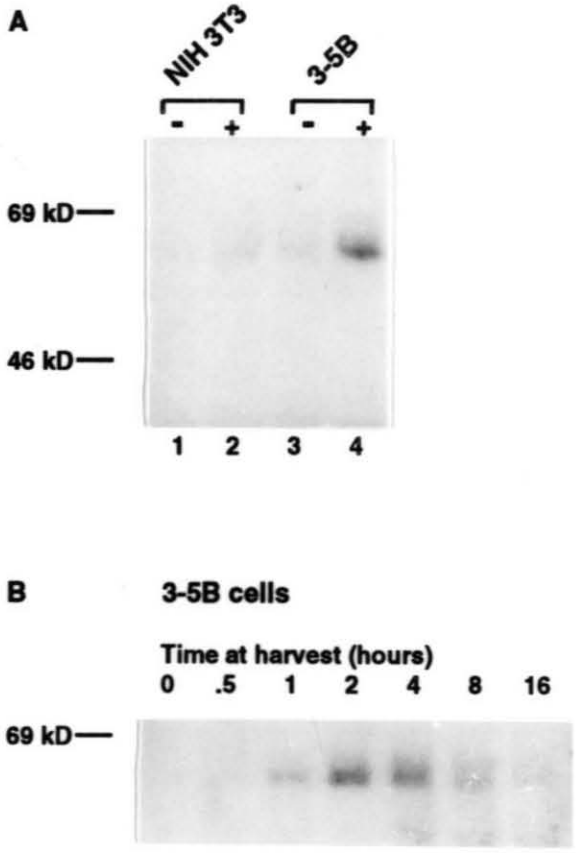


Figure 3. Emergence from growth arrest stimulated by conditional myc expression. **A.** NIH 3T3 and conditional myc23c expressing 3-5B cells were seeded and brought to growth arrest in defined medium as indicated in materials and methods. Arrested cells were mock or zinc stimulated; zinc doses are indicated at the bottom of each column. **B.** Similar experiments comparing parental Balb/c 3T3 cells with the c-myc estrogen receptor fusion protein (myc ER) expressing cell line MER #6. Cells were seeded and brought to growth arrest in medium containing 5% calf serum as indicated in materials and methods. Arrested cells were mock or β-estradiol (100 nM) stimulated as indicated at the bottom of each column. Cells were labelled with BrdU from 14 to 24 hours following stimulation. Fixation and immunoperoxidase staining were as in materials and methods. Approximately 400 cells were counted for each zinc dose; labelling index is the number of BrdU stained cells divided by the total number of cells.

Figure 4. Western analysis of c-myc protein expression in NIH 3T3 and myc23c expressing 3-5B cells. **A.** (-): cells were mock induced. (+): 3T3 cells were induced by adjusting the medium to 5% fetal bovine serum. 3-5B cells were induced by adjusting the medium to 20 μM $ZnSO_4$. All cells were harvested after 4 hours of induction. **B.** 3-5B cells were harvested over a 16 hour timecourse following induction. Cells in the "0" timepoint were harvested 4 hours after mock induction. The other samples were harvested at the times indicated at the tops of the lanes. Cells were seeded, brought to growth arrest, induced, harvested, and blotted as indicated in materials and methods. The labelling antibody was ^{125}I conjugated anti c-myc monoclonal H60C37.

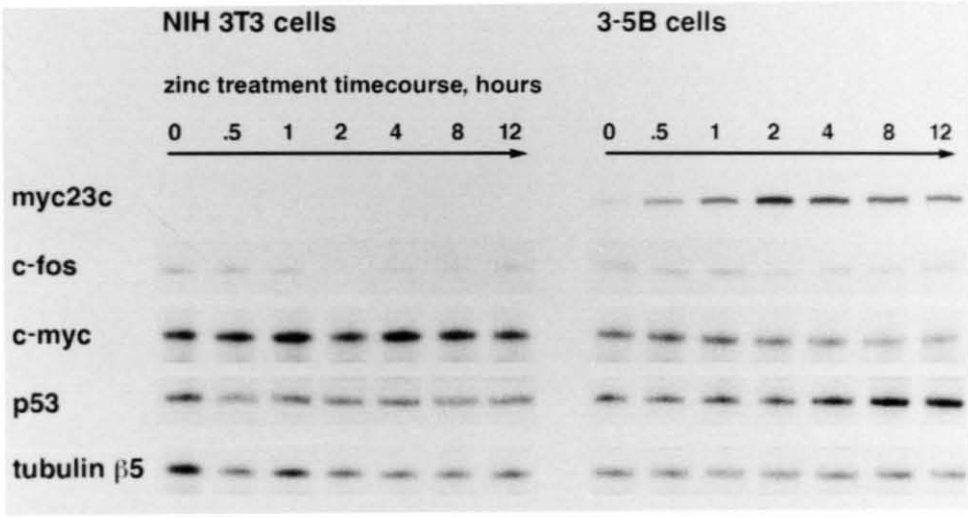

A
NIH 3T3
3-5B
- + - +
69 kD
46 kD
1 2 3 4
B
3-5B cells
Time at harvest (hours)
0 .5 1 2 4 8 16
69 kD

Balb/c 3T3 cells expressing a c-myc estrogen receptor fusion protein reenter the cell cycle following hormonal activation of the fusion protein (Eilers et al., 1991; see also figure 3).

As an initial calibration of RNA level responses to conditional c-myc in our system, we examined the expression of three already known G_1-regulated transcripts (the immediate early genes c-fos and endogenous c-myc, and the mid-SR gene p53 tumor antigen; tubulin β5 serves as an unaffected control) (Figure 5). In both parental 3T3 cells and myc23c expressing 3-5B cells, c-fos RNA was present at very low levels and did not vary significantly with c-myc expression. In accord with other work (Eilers et al., 1991), this result demonstrated with high sensitivity that the immediate early burst of c-fos RNA synthesis following serum stimulation is bypassed in a c-myc stimulated emergence from G_0. As in other myc transgene expressing cell lines, endogenous c-myc RNA levels were chronically repressed in 3-5B cells relative to 3T3 cells (Cleveland et al., 1988; Grignani et al., 1990; Penn et al., 1990); this was probably a consequence of elevated basal expression of transgene c-myc in 3-5B cells. In addition, the endogenous c-myc transcript was further repressed over the timecourse of the myc transgene induction. That myc itself shows different regulation in response to full serum stimulation (up-regulation at mid-G_1 relative to G_0) compared to the response to ectopic transgene myc (down-regulation relative to G_0), raises the possibility that other genes from the mid-SR panel might display disparate responses in our mitogen- versus myc-driven $G_0 \rightarrow S$ experimental regimes.

p53 transcript levels remained steady in control 3T3 cells treated with zinc, but they were induced 2- to 3-fold between 4 and 12 hours after zinc addition to 3-5B cells, consistent with its timecourse of up-regulation in response to serum stimulation (Reich and Levine 1984). This identifies p53 as a target, direct or indirect, of myc in this experimental regime. This result is consistent with previous observations that p53 is among the proteins that are overexpressed in

Figure 5. Comparison of selected mRNA expression in NIH 3T3 and myc23c expressing 3-5B cells. NIH 3T3 and 3-5B cells were seeded and brought to growth arrest in defined medium as indicated in materials and methods. Arrested cells were either mock stimulated or treated with 20 μM zinc. RNAs generated from mock and zinc treated cells were analyzed by ribonuclease protection. Timecourse: 0, mock stimulated cells; .5→12, cells treated with zinc for 0.5 to 12 hours. Please note that the ratio of the sensitivity of the c-myc to myc23c protections was 8:1; the sensitiviy of the c-fos protection was equivalent to that of the c-myc protection, just sufficient to detect basal c-fos expression.


NIH 3T3 cells
3-5B cells
zinc treatment timecourse, hours
0 .5 1 2 4 8 12
0 .5 1 2 4 8 12
myc23c
c-fos
c-myc
p53
tubulin β5

cells whose growth characteristics have been altered by chronic c-myc overexpression (Shiroki et al., 1986), although it has been difficult in such systems to discriminate obligate responses to myc from various indirect effects. Nevertheless, among known candidates for direct regulation by myc, p53 presently stands out as a good candidate for direct interaction with myc because its expression in transfection assays is dependent on an E-box sequence element reminiscent of the *in vitro*-defined myc binding site (Ronen et al., 1991). The presence of such an E-box is only suggestive of a direct myc interaction, since this E-box sequence is also a target for strong *in vitro* binding by MLTF/USF (Carthew et al., 1985; Gregor et al., 1990), max/myn homodimers (Blackwood and Eisenman 1990; Prendergast et al., 1991) and TFII-i (Roy et al., 1991). However, the observed response to conditional myc shown here provides additional support for a direct regulatory interaction between myc and p53. The data also suggest that the magnitude of myc response expected from unknown target genes in the mid-SR panel may be quantitatively modest, as is the case for p53.

Identification of myc regulated transcripts among the panel of mid-SR regulated clones. The metallothionein promoted c-myc gene expressed by 3-5B cells included a glucocorticoid regulated enhancer derived from MMTV which was positioned upstream of the promoter and metal regulatory elements from the murine metallothionein gene. We had originally planned to use this second, metal-independent myc induction system to control for possible phenotypic effects secondary to metal treatment. This would have substantially improved confidence by showing a myc response via two independent conditional expression systems. Unfortunately, regulation of the myc23c transgene by glucocorticoids proved weak and inconsistent in this cellular environment and was not pursued further. Instead, we substituted the cell line MER #6 which expresses a c-myc estrogen receptor fusion protein (myc-ER) in the Balb/c 3T3 cell background (Eilers et al., 1989; Eilers et al., 1991) (a

kind gift from Sabine Schirm and J. Michael Bishop). The format for characterizing the effect of myc on the expression of the individual members of the mid-SR collection was straightforward. NIH 3T3 and 3-5B cells were brought to growth arrest at low density in defined medium and either mock or zinc stimulated; similarly, Balb/c 3T3 and MER #6 cells were brought to growth arrest at low density in defined medium and either mock or estrogen stimulated. The mock stimulation conditions included medium changes, inducing agent solvents and manipulations parallel to metal or hormone inductions. RNAs were harvested 8 hours after mock, zinc, or estrogen treatment. Transcript levels were measured by ribonuclease protection and quantitated on a phosphorimager.

Table 1 and figure 6 present the effect of conditional myc expression on the mid-SR regulated cDNAs. Each data point from the NIH^3-5B system results from a single experiment; results from the Balb/c^MER #6 system are averaged from 3 successive experiments. To more simply visualize myc response in the two conditional myc systems, figure 6 presents data from Table 1 in a scatter diagram format. In figure 6A, each cDNA is represented as a point where fold-regulation by the c-myc estrogen receptor fusion protein in MER #6 cells is the X-value and fold-regulation by c-myc in 3-5B cells is the Y-value; for those transcripts that lie away from the central cluster, the X-error bars delimit +/- one standard deviation as derived from three independent experiments with the Balb/c and MER #6 cells. While most of the transcripts cluster around (1,1), indicative of no significant response to elevated myc expression, I-8-10B [lactate dehydrogenase], I-8-36 [ornithine decarboxylase], I-8-29 [unknown], and I-8-30B [unknown] all lie in the upper right quadrant, indicative of induction by myc. To better visualize myc-driven repression, the data were replotted in figure 6B as the reciprocal of the values in 6A. In this analysis, I-8-1 [tenascin] and I-8-41 [heat stable antigen] lie in the upper right quadrant, indicative of repression by myc. Of the 18 mid-SR

Table 1A. cDNA clones subject to serum regulation with mid-SR kinetics

	Cell system:	NIH 3T3	3-5B	MER #6
	Induction:	by serum[1] at 8 hours	by conditional myc[2] expression at 8 hours	
clone #	clone identity			
I-8-1	tenascin	110	0.69	0.33 +/- .2
I-8-6	MGSA[3]	8.6	1.0	0.66 +/- .1
I-8-9	not identified	3.6	1.0	0.84 +/- .3
I-8-10A	mouse T1 protein	26	0.9	not detected
I-8-10B	lactate dehydrogenase	8.9	2.21	2.0 +/- .1
I-8-16	A-X actin	4.6	0.6	0.88 +/- .1
I-8-23	not identified	3.2	not done	0.96 +/- .3
I-8-29	not identified[4]	5.1	1.73	5.0 +/- .6
I-8-30B	not identified	2.0	2.5	1.4 +/- .2
I-8-32	liver thiol transferase	7.0	1.1	0.86 +/- .1
I-8-33	G-3-P dehydrogenase[5]	2.9	1.2	1.0 +/- .1
I-8-34	not identified	28	0.89	1.3 +/- .7
I-8-36	ornithine decarboxylase	2.6	3.96	1.7 +/- .2
I-8-41	heat stable antigen	4.6	not detected	0.5 +/- .1
I-8-47	not identified	70	1.49	1.4 +/- .2
	Repression:[6]	by serum at 12 hours	by conditional myc expression at 8 hours	
R-8-1	Alpha-1 (3) collagen	50	1.7	1.0 +/- .1
R-8-9	elongation factor 1	2.4	not done	1.1 +/- .2
R-8-26	Alpha-1 (6) collagen	2.7	not done	1.1 +/- .1

Table 1B. cDNA clones subject to serum regulation with mid-SR kinetics

Cell system:	NIH 3T3	3-5B	MER #6
Induction:	by serum at 8 hours	by conditional myc expression at 8 hours	
alpha prothymosin	1.7	3.3	1.2 +/- .1
c-myc	15	0.36	0.76 +/- .4
G1 cyclin D2	3.1	1.6	1.5 +/- .2
max/myn	2.4	2.1	1.2 +/- .3
p53 tumor antigen	3.8	2.3	1.2 +/- .2
c-fos	0.83	0.96	not done
c-jun	0.80	0.79	not done
fra-1	8.5	1.0	not done
HSP 70 MHS 214	3.0	1.2	not done
Id	31	0.89	not done

1. Induction by serum: transcript prevalence in serum stimulated cells divided by transcript prevalence in growth arrested cells.

2. Induction by conditional myc expression:

$$\frac{\text{fold change upon induction in conditional cell line}}{\text{fold change upon induction in control 3T3 cell line}}$$

3. MGSA: melanoma growth stimulatory activity

4. Shares some sequence homology with methylenetetrahydrofolate synthetase.

5. G-3-P dehydrogenase: glyceraldehyde-3-phosphate dehydrogenase.

6. Repression values are the reciprocal of induction values

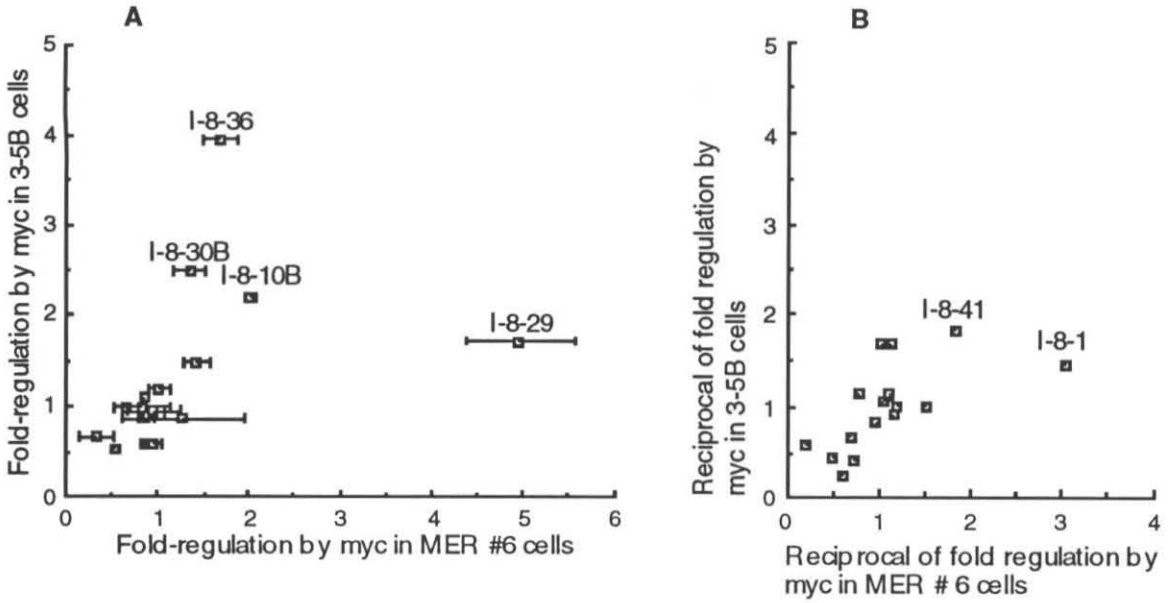


Figure 6. Scatter diagram analysis of mid-SR regulated transcript expression in conditional myc expressing 3-5B and MER #6 cells. Data are from table 1. **A.** Data for each transcript are plotted on a scatter diagram. The X-value for each point is:

$$\frac{\text{fold change upon induction in MER \#6 cells}}{\text{fold change upon induction in Balb/c 3T3 cells}}$$

The Y-value for each point is:

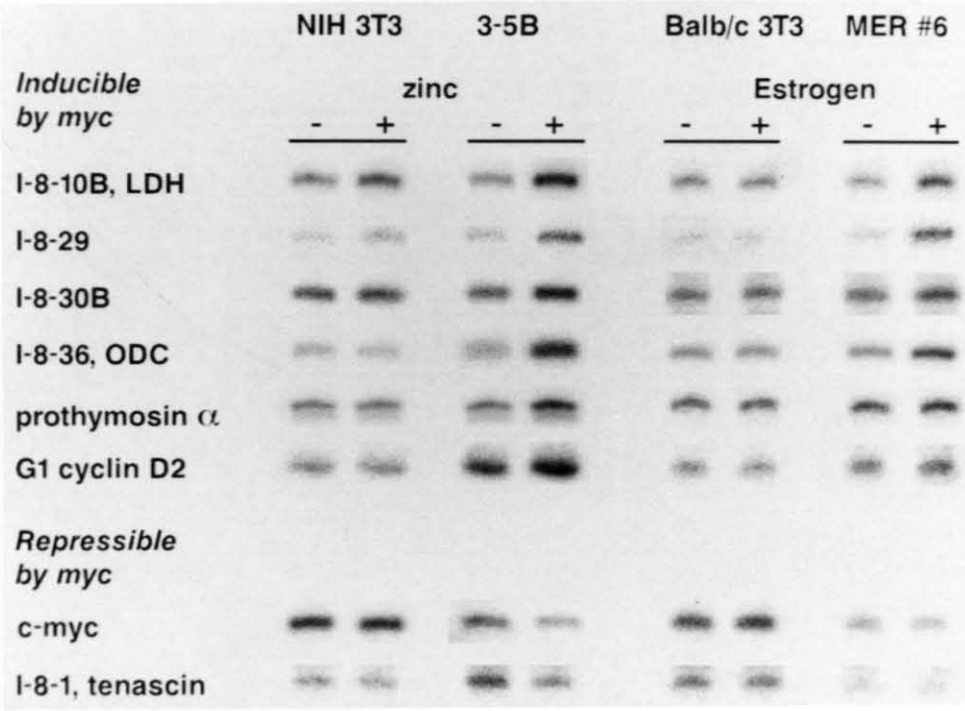
$$\frac{\text{fold change upon induction in 3-5B cells}}{\text{fold change upon induction in control NIH 3T3 cells}}$$

The Balb/c^MER #6 experiment was iterated 3 times; the plotted data are average values and the X-error bars delimit +/- one standard deviation for these data. Clone numbers are provided for those transcripts that were clearly upregulated by conditional myc expression. **B.** Data for each transcript are replotted on a scatter diagram where the X- and Y-values are the reciprocals of the values in part A. Clone numbers are provided for those transcripts that were clearly downregulated by conditional myc expression.

regulated mRNAs, 5 (lactate dehydrogenase, ornithine decarboxylase, tenascin, I-8-29, and I-8-30B) are induced or repressed more than 2-fold following stimulation of myc expression in one of the two systems; in each case, that same mRNA is regulated in the same manner, though not necessarily to the same degree, in the alternate myc expressing cell system. I-8-47 was about 1.5-fold induced following expression of myc in both cell systems. Expression from the immediate early serum response genes c-fos, c-jun, and c-myc and the previously identified myc-responsive gene prothymosin α were also measured in this experiment. Myc behaved essentially as before in 3-5B cells (Figure 5), while in MER #6 cells myc was repressed to very low levels, presumably due to basal activity from the myc-ER fusion protein (Eilers et al., 1989). Neither c-fos nor c-jun responded to myc expression in either system. Prothymosin α was induced in 3-5B cells, confirming in this NIH 3T3 cellular background the original observation myc of inducibility reported in MycER expressing Rat 1 cells (Eilers et al., 1991). Prothymosin α was not, however, induced above background in MER #6 cells, in agreement with prior studies (Hevezi and Bishop, personal communication). For the key myc responses reported here, primary RNA level data are presented in Figure 7.

Viewed as a whole, the myc and serum screening results suggest that roughly one third of mid-SR induced or repressed cDNAs can also be regulated, at least in part, by c-myc, although the data do not reveal whether myc-regulation is direct or indirect. The screening results also revealed that among the mid-SR genes that were most profoundly regulated in response to serum, there are several that were unaffected or even affected inversely by myc. As discussed below, this suggests that a myc-driven $G_0 \rightarrow S$ progression differs qualitatively from a parallel, serum-driven progression in potentially interesting and previously undetected ways.

Figure 7. Expression of selected myc regulated transcripts. NIH 3T3, myc23c expressing 3-5B, Balb/c 3T3, and mycER expressing MER #6 cells were seeded and brought to growth arrest in defined medium as indicated in materials and methods. NIH 3T3 and 3-5B cells were mock or zinc treated; Balb/c and MER #6 cells were β-estradiol treated. Cells were harvested after 8 hours of stimulation. RNAs were analyzed by ribonuclease protection. The induced zinc condition was 20 μM for all transcripts except for I-8-29, for which it was 15 μM. The induced β-estradiol condition was 100 nM in all cases. Abbreviations: LDH, lactate dehydrogenase; ODC, ornithine decarboxylase.


NIH 3T3
3-5B
Balb/c 3T3
MER #6
Inducible by myc
zinc
Estrogen
- + - + - + - +
I-8-10B, LDH
I-8-29
I-8-30B
I-8-36, ODC
prothymosin α
G1 cyclin D2
Repressible by myc
c-myc
I-8-1, tenascin

Discussion

The emergence of cells from a quiescent G_0 arrested state into the cell cycle is a multistep process that begins with the immediate early response to mitogens and extends into the specialized G_1 phase that follows (for review, see Pardee 1989). The immediate early response occurs in the first hour of this mitogen initiated G_1 and is dominated by new transcription of genes encoding transcriptional regulators including c-fos, jun and c-myc. The distinction between this G_1 and that of exponentially growing cells is already evident at this early time point, since myc and fos are not detectably up-regulated at the Mitosis/G_1 boundary in cycling cells but are instead expressed at a much lower and constant level throughout the cycle (Thompson et al., 1985; Blackwood et al., 1992). Identification of many of the immediate early genes as transcriptional regulators has raised the question of what their regulatory targets are and how those targets contribute to progression toward S-phase. In this work we have cloned cDNA copies of messenger RNAs that are either up- or down-regulated at a later time in the serum initiated G_1 cascade with the expectation that these will include direct and indirect targets of immediate early regulators. In particular, for the case of myc regulated RNAs, the pattern of expression for myc and max/myn suggested that the potential c-myc:{max/myn} transcription factor should be prominent during the mid-G_1 period following serum stimulation of arrested cells. From 18 different mid-SR class cDNAs isolated, a subset that respond in a myc-driven progression through G_0 were selected.

Serum-driven and myc-driven emergence from G_0 are qualitatively and quantitatively different. Following stimulation of myc expression in quiescent cells, both 3-5B and MER #6 cells emerge from growth arrest, progress through G_1, and enter S-phase. However, in these systems only a fraction (usually 40-50%) of the arrested cells

ultimately enter S-phase. In contrast, >95% of these same cells will enter S-phase following serum stimulation, which suggests that there is at least a significant quantitative difference in the responsiveness of cells to myc versus serum; a substantial qualitative difference might also exist. These are not mutually exclusive possibilities, but our results emphasize to an unanticipated extent the existence of substantial qualitative differences.

Comparison of the serum response with myc response first shows that many genes strongly regulated by serum do not respond detectably to myc. This result is consistent with expectations based on earlier studies (Eilers et al., 1991; Yang et al., 1991). Thus, some prominent members of the immediate early serum response are not induced by ectopic myc; rather, in a myc driven emergence from G_0 regulators such as fos are apparently bypassed. It is therefore reasonable to expect that there will be some genes whose induction in response to serum depends heavily or entirely on the bypassed regulators. Even the limited panel of mid-SR genes presented here includes a dozen that are good candidates for regulation by non-myc immediate early regulators. At the cellular level, we can conclude that these myc-insensitive mid-SR products are not essential for progression to S-phase, although they may account for differences in the character and quantitative success of the full mitogen response compared with the myc response. Viewed from a different perspective, it is even possible that some mid-SR genes that are absent from the myc response class contribute to myc induced progression through their absence by virtue of normally limiting progression toward S-phase in a full serum response. Finally, the complex differences between the myc-responsive set and the larger serum responsive set highlight a practical point. It is possible that there exist genes that are regulated by myc in the absence of serum that are <u>not</u> regulated at all by serum alone; in our two step screen for myc response these would have been absent from the starting pool. If these exist, a direct ± myc subtractive screen will be needed to find them.

A second and extreme sort of divergence between serum and myc driven G_1 phases is revealed in the cases of myc itself and tenascin [I-8-1], an extracellular matrix protein. Both were substantially up-regulated in mid-G_1 by serum, while ectopic myc expression resulted in modest but reproducible down-regulation. This divergence in regulatory pattern clearly shows that the myc stimulated pathway is not a simple subset of the full serum stimulated response. It will now be crucial to determine the mechanistic basis for each type of response.

Although there were major qualitative and quantitative differences in myc response versus serum response for many of the mid-SR cDNAs, there were also some that responded similarly. Among these were lactate dehydrogenase [I-8-10B], I-8-29, ornithine decarboxylase [I-8-36], and p53 tumor antigen. These similarly affected products are good candidates for proteins obligatorily involved in the regulation or execution of any successful progression from G_0 to S-phase.

Some genes are members of both immediate early and mid-SR groups. A good example of the anticipated and probably complex overlap between immediate early and mid-SR genes is ornithine decarboxylase (I-8-36), which is up-regulated in a c-myc driven cycle and in serum driven re-entry, but is also an immediate early gene (induced rapidly by serum in a protein synthesis independent fashion) by virtue of the presence of a functional CRE related element within its promoter (Palvimo et al., 1991). That such overlap might be observed was foreshadowed by the fact that c-myc is an immediate early regulator, but also continues to be expressed in mid-G_1. Indeed, we have found that ectopic myc expression in 3-5B cells and similar lines must be expressed throughout G_1 in order to support entry into S-phase (chapter 1), showing that it is also functional beyond the immediate early period.

Other myc targets. Several prior studies have been directed at identification of myc regulatory targets, and some of these targets were

screened in the course of this work for responsiveness in the $G_1 \rightarrow$ S-phase inducible system. Chronic overexpression of myc leads to transcriptional repression of c-myc and to repression of MHC class I transcripts (Bernards et al., 1986; Cleveland et al., 1988; Versteeg et al., 1988; Grignani et al., 1990; Penn et al., 1990). We found here that myc autorepression is also observed upon short-term myc overexpression. In constitutive myc overexpressing lines, Prendergast and Cole were able to identify two myc regulated clones, MR1 (plasminogen activator inhibitor 1), and MR2. At the mechanistic level, their data suggested that regulation was post-transcriptional in both cases (Prendergast et al., 1990). Yang et al. identified three collagen transcripts transcriptionally repressed by c-myc overexpression, and Eilers et al. found that prothymosin α is transcriptionally induced by a c-myc estrogen receptor fusion protein in a Rat-1 cell background, though not in a Balb/c 3T3 cell background (Eilers et al., 1991; Yang et al., 1991, Hevezi and Bishop, personal communication). Here we also observed its failure to respond in the Balb/c based MER#6 cells, but found it to be myc-responsive in the NIH 3T3 based 3-5B cells. Except for a collagen family member, our mid-SR screens did not identify any of the previously identified targets. This is expected because the screening was not extensive enough to be saturating, even though the libraries themselves are comprehensive. Additional rounds of screening would be expected to yield substantial numbers of new mid-SR and myc responsive clones.

Differences and similarities in myc response in MER#6 and 3-5B myc induction systems. For the mid-G1 regulated genes obtained in our cloning effort, myc regulated expression in the NIH^3-5B and Balb/c^MER #6 systems corroborated nicely; there was a 1:1 correspondence between down-regulated, unregulated, and up-regulated transcripts. This result lends confidence that the observed responses are indeed attributable to myc. The correspondence may also reflect the

fact that we intentionally tried to filter out cell line specific differences by screening the Balb/c 3T3 based libraries with NIH 3T3 serum response probes. In contrast, we did find differences in the responses of myc target genes that were selected for screening in the conditional myc expression systems independent of their appearance in the mid-SR panel. These included prothymosin α, G1 cyclin D2, p53 tumor antigen, and max transcripts, all of which were all up-regulated in 3-5B but not MER #6 cells. For most of these we are not yet certain whether they are legitimate myc targets, and if they are, whether the observed differences in myc response stem from a functional distinction between c-myc and the c-myc estrogen receptor fusion protein or reflect differences between NIH 3T3 and Balb/c 3T3 cells. In the specific case of prothymosin α, its identity as a gene induced by myc in some, but not all, cell types is strengthened by the fact that it was myc responsive in Rat-1 cells carrying the mycER fusion protein (Eilers et al., 1991) and in the NIH 3T3 based 3-5B cells.

Eventually, studies of myc and its regulatory targets should be able to explain both the normal function and immortalizing potential of c-myc in terms of the regulatory pathways in which it participates and the functions of the gene products that it regulates. At this very early point, the identities of some of the cDNAs isolated are interesting. In the direction of promoting cell cycle progression, ornithine decarboxylase catalyzes the first step in polyamine synthesis; its activity is a necessary correlate of DNA synthesis. Beyond the ability of mammalian G1 cyclins to form complexes with p34 cdc2 and cdk protein kinases and complement yeast G_1 cyclin deficiencies (Koff et al., 1991; Lew et al., 1991; Matsushime et al., 1991), their functions have not yet been worked out. However G1 cyclin D2, which lies downstream of the CSF-1 receptor (Matsushime et al., 1991) is up-regulated by c-myc; furthermore, a requirement for proper c-myc expression can be dissected away from stimulation of c-jun and c-fos expression via CSF-1 receptor mutants (Roussel et al., 1991).

Mechanisms of myc responsiveness. At the level of characterization presented here, we cannot know how many regulatory steps occur between accumulation of c-myc protein and up- or down-regulation of a particular myc responsive mRNA. Also, we do not know the mechanism(s) responsible; they may be either transcriptional, post-transcriptional or a combination of both. However, it is interesting to note that the promoters of lactate dehydrogenase, ornithine decarboxylase, and p53 tumor antigen all contain potential myc binding sites as previously defined by *in vitro* target site selection experiments (Blackwell et al., 1990; Halazonitas and Kandil 1991; Kerkhoff et al., 1991). We have recently shown that the candidate site within LDH is a good *in vitro* binding site for myc:max/myn heteromers or max/myn homo-oligomers (Appendix A). Also, a potential myc binding site within the mouse p53 promoter, located just downstream of the cap site, is a positive regulator of p53 expression (Ronen et al., 1991). While other transcription factors, notably USF/MLTF and TFII-i (Carthew et al., 1985; Gregor et al., 1990; Roy et al., 1991), can recognize the same core E-box sequence, these genes may in fact be directly regulated by c-myc/max in its role as a transcription factor or by max/myn homo-oligomers. It will now be crucial to understand in detail the balance between levels of myc and max/myn hetero- and homo-oligomers and how these levels correlate with the responses to myc and serum reported here. A relevant observation from this work was that induction of ectopic myc in 3-5B cells results in up-regulation of max/myn RNA, so that active myc:max/myn heteromers are probably able to form. However, if max/myn levels exceed available myc protein, max/myn homodimers may compete successfully for DNA binding sites with myc:max heteromers (Blackwood and Eisenman 1990; Prendergast et al., 1991; Makela et al., 1992) to antagonize the effect of a myc containing transcription factor.

References

Allen-Hoffman, B. L., Crankshaw, C. L., and Mosher, D. F. (1988). Transforming growth factor B increases cell surface binding and assembly of exogenous (plasma) fibronectin by normal human fibroblasts. Molecular and Cellular Biology **8**, 4234-4242.

Almendral, J. M., Sommer, D., MacDonald-Bravo, H., Burckhardt, J., Perera, J., and Bravo, R. (1988). Complexity of the early genetic response to growth factors in mouse fibroblasts. Molecular and Cellular Biology **8**, 2140-2148.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. Journal of Molecular Biology **215**, 403-410.

Armelin, H. A. M., Armelin, C. S., Kelly, K., Stewart, T., Leder, P., Cochran, B. H., and Stiles, C. D. (1984). Functional role for c-myc in mitogenic response to platelet-derived growth factor. Nature **310**, 655-660.

Benezra, R., Davis, R. L., Lockson, D., Turner, D. L., and Weintraub, H. (1990). The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. Cell **61**,

Bernard, O., Cory, S., Gerondakis, S., Webb, E., and Adams, J. M. (1983). Sequence of the murine and human cellular myc oncogenes and two modes of myc transcription resulting from chromosome translocation in B lymphoid tumours. European Molecular Biology Organization Journal **2**, 2375-2383.

Bernards, R., Dessain, S., and Weinberg, R. (1986). N-myc amplification causes down-modulation of MHC class I antigen expression in neuroblastoma. Cell **47**, 667-674.

Blackwell, T. K., Kretzner, L., Blackwood, E. M., Eisenman, R. N., and Weintraub, H. (1990). Sequence-specific DNA binding by the c-myc protein. Science **250**, 1149-1151.

Blackwood, E. M., and Eisenman, R. N. (1990). Max: A helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with myc. Science **251**, 1211-1217.

Blackwood, E. M., Luscher, B., and Eisenman, R. N. (1992). Myc and max associate in vivo. Genes and Development **6**(1), 71-80.

Bond, V. C., and Wold, B. (1987). Poly-l-ornithine-mediated transformation of mammalian cells. Molecular and Cellular Biology **7**, 2286-2293.

Bravo, R. (1990). Growth factor-responsive genes in fibroblasts. Cell growth and differentiation **1**, 305-309.

Carthew, R. W., Chodosh, L. A., and Sharp, P. A. (1985). An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. Cell **43**, 439-448.

Cavalieri, F., and Goldfarb, M. (1987). Growth-factor-deprived BALB/c 3T3 murine fibroblasts can enter the S phase after induction of c-myc gene expression. Molecular and cellular biology **7**, 3554-3560.

Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162**, 156-159.

Cleveland, J. L., Huleihel, M., Bressler, P., Siebenlist, U., Akiyama, L., Eisenman, R. N., and Rapp, U. R. (1988). Negative regulation of c-myc transcription involves myc family proteins. Oncogene Research **3**, 357-375.

Eilers, M., Picard, D., Yamamoto, K. R., and Bishop, J. M. (1989). Chimaeras of Myc oncoprotein and steroid receptors cause hormone-depndent transfromation of cells. Nature **340**, 66-68.

Eilers, M., Schirm, S., and Bishop, J. M. (1991). The myc protein activates transcription of the alpha-prothymosin gene. European Molecular Biology Organization Journal **10**(1), 133-141.

Gregor, P. D., Sawadogo, M., and Roeder, R. G. (1990). The adenovirus major late transcription factor USF is a member of the helix-loop-helix group of regulatory proteins and binds DNA as a dimer. Genes and Development **4**, 1730-1740.

Grignani, F., Lombardi, L., Inghirami, G., Sternas, L., Cechova, K., and Dalla-Favera, R. (1990). Negative autoregulation of c-myc gene expression is inactivated in transformed cells. European Molecular Biology Organization Journal **9**(12), 3913-3922.

Halazonitas, T. D., and Kandil, A. N. (1991). Determination of the c-myc DNA-binding site. Proceedings of the National Academy of Sciences USA **88**, 6162-6166.

Jendrisak, J., Young, R. A., and Engel, J. D. (1987). Cloning of cDNA into lambda gt10 and lambda gt11. Guide to molecular cloning techniques. San Diego, Academic Press, Inc. 359-371.

Kerkhoff, E., Bister, K., and Klempnauer, K. H. (1991). Sequence-specific DNA binding by myc proteins. Proceedings of the National Academy of Sciences USA **88**, 4323-4327.

Koff, A., Cross, F., Fisher, A., Schumacher, J., Leguellec, K., Phillipe, M., and Roberts, J. M. (1991). Human cyclin E, and new cyclin that interacts with two members of the cdc2 family. cell **66**(6), 1217-1228.

Lau, L. F., and Nathans, D. (1987). Expression of a set of growth-related immediate early genes in Balb/c 3T3 cells: coordinate regulation with c-fos or c-myc. Proceedings of the National Academy of Sciences USA **84**, 1182-1186.

Lew, D. J., Dulic, V., and Reed, S. I. (1991). Isolation of three novel human cyclins by rescue of G1 cyclin (cln) function in yeast. Cell **66**(6), 1197-1206.

Lewis, S. A., Lee, M. G. S., and Cowan, N. J. (1985). Five mouse tubulin isotypes and their regulated expression during development. Journal of Cell Biology **101**, 305-321.

Lowe, D. G., and Moran, L. A. (1986). Molecular cloning and analysis of DNA complementary to three mouse M_r = 68,000 heat shock protein mRNAs. The Journal of Biological Chemistry **261**(5), 2102-2112.

Makela, T. P., Koskinen, P. J., Vastrik, I., and Alitalo, K. (1992). Alternative forms of max as enhancers or suppressors of myc-ras cotransformation. Science **256**, 373-376.

Maniatis, T., Fritsch, E. F., and Sambrook, J., Ed. (1982). Molecular Cloning. Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory.

Matsushime, H., Roussel, M. F., Ashmun, R. A., and Sherr, C. J. (1991). Colony-stimulating factor 1 regulates novel cyclins during the G1 phase of the cell cycle. Cell **65**, 701-713.

Mirkes, P. E. (1985). Simultaneous banding of rat embryo DNA, RNA, and protein in cesium trifluoroacetate gradients. Analytical Biochemistry **148**, 376-383.

Mohn, K. L., Laz, T. M., Hsu, J.-C., Melby, A. E., Bravo, R., and Taub, R. (1991). The immediate-early growth response in regenerating liver and insulin-stimulated H-35 cells: comparison with serum stimulated 3T3 cells and identification of 41 novel immediate-early genes. Molecular and Cellular Biology **11**(1), 381-390.

Mueller, P. R., Salser, S. J., and Wold, B. (1988). Constitutive and metal-inducible protein:DNA interactions at the mouse metallothionein I promoter examined by in vivo and in vitro fooptrinting. Genes and Development **2**, 412-427.

Muller, R., Bravo, R., Burckhardt, J., and Curran, T. (1984). Induction of c-fos gene and protein by growth factors precedes activation of c-myc. Nature **312**, 716-720.

Palazzolo, M. J., Hamilton, B. A., Ding, D., Martin, C. H., Mead, D. A., Mierendorf, R. C., Raghavan, K. V., Meyerowitz, E. M., and Lipshitz, H. D. (1990). Phage lambda cDNA cloning vectors for subtractive hybridization, fusion protein synthesis and Cre-loxP automatic plasmid subcloning. Gene **88**, 25-36.

Palvimo, J. J., Eisenberg, L. M., and Janne, O. A. (1991). Protein-DNA interactions in the cAMP responsive promoter region of the murine ornithine decarboxylase gene. Nucleic Acids Research **19**(14), 3921-3927.

Pardee, A. B. (1989). G1 events and regulation of cell proliferation. Science **246**, 603-608.

Payvar, F., DeFranco, D., Firestone, G. L., Edgar, B., Wrange, O., Okret, S., Gustafsson, J. A., and Yamamoto, K. R. (1983). Sequence-specific binding of glucocorticoid receptor to MTV DNA at sites within and upstream of the transcribed region. Cell **35**, 381-392.

Penn, L. J. Z., Brooks, M. W., Laufer, E. M., and Land, H. (1990). Negative autoregulation of c-myc transcription. European Molecular Biology Organization Journal **9**(4), 1113-1121.

Prendergast, G. C., Diamond, L. E., Dahl, D., and Cole, M. D. (1990). the c-myc regulated gene mr1 encodes plasminogen activator inhibitor-1. Molecular and Cellular Biology **10**, 1265-1269.

Prendergast, G. C., Lawe, D., and Ziff, E. B. (1991). Association of Myn, the murine homolog of max, with c-myc stimulates methylation-sensitive DNA-binding and ras cotransformation. Cell **65**(3), 395-407.

Prendergast, G. C., and Ziff, E. B. (1991). Methylation-sensitive sequence-specific DNA binding by the c-myc basic region. Science **251**, 186-189.

Reich, N., and Levine, A. J. (1984). Growth regulation of a cellular tumour antigen, p53, in nontransformed cells. Nature **308**, 199-201.

Ronen, D., Rotter, V., and Reisman, D. (1991). Expression from the murine p53 promoter is mediated by a factor binding to a downstream helix-loop-helix recognition motif. Proceedings of the National Academy of Sciences USA **88**, 4128-4132.

Roussel, M. F., Cleveland, J. L., Shurtleff, S. A., and Sherr, C. J. (1991). Myc rescue of a mutant CSF-1 receptor impaired in mitogenic signalling. Nature **353**(6342), 361-363.

Roy, A. L., Pognonec, M. M. P., and Roeder, R. G. (1991). Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF. Nature **354**, 245-248.

Sargent, T. D. (1987). Isolation of differentially expressed genes. Guide to molecular cloning techniques. San Diego, Academic Press, Inc. 423-432.

Shiroki, K., Segawa, K., Koita, Y., and Shibuya, M. (1986). Neoplastic transformation of rat 3Y1 cells by a transcriptionally activated human c-myc gene and stabilization of p53 cellular tumor antigen in the transformed cells. Molecular and Cellular Biology **6**(12), 4379-4386.

Tan, T., Wallis, J., and Levine, A. J. (1986). Identification of the p53 Protein Domain Involved in Formation of the Simian Virus 40 Large T-Antigen-p53 Protein Complex. Journal of Virology **59**, 574-583.

Thompson, C. B., Challoner, P. B., Neiman, P. E., and Groudine, M. (1985). Levels of c-myc oncogene mRNA are invariant throughout the cell cycle. Nature **314**, 363-369.

Van Beveren, C., Straaten, F. v., Curran, T., Muller, R., and Verma, I. M. (1983). Analysis of FBJ-MuSV provirus and c-fos (mouse) gene reveals that viral and cellular fos gene products have different carboxy terminii. Cell **32**, 1241-1255.

Versteeg, R., Noordermeer, I., Kruse-Wolters, M., Ruiter, D., and Schrier, P. (1988). c-myc down-regulates class I HLA expression in human melanomas. European Molecular Biology Organization Journal **7**, 1023-1029.

Wenzel, A., Cziepluch, C., Hamann, U., Schurmann, J., and Schwab, M. (1991). The N-myc oncoprotein is associated in vivo with the phosphoprotein max(p20/22) in human neuroblastoma cells. European Molecular Biology Organization Journal **10**(12), 3703-3712.

Wigler, M., Sweet, R., Sim, G. K., Wood, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S., and Axel, R. (1979). Transformation of mammalian cells with genes from procaryotes and eucaryotes. Cell **16**, 777-785.

Yang, B. S., Geddes, T. J., Pogulis, R. J., de Crombrugghe, B., and Freytag, S. O. (1991). Transcriptional suppression of cellular gene expression by c-myc. Molecular and Cellular Biology **11**(4), 2291-2295.

Zhan, X., and Goldfarb, M. (1986). Growth factor requirements of oncogene-transformed NIH 3T3 and BALB/c 3T3 cells cultured in defined media. Molecular and Cellular Biology **6**, 3541-3544.

Zinn, K., Dimaio, D., and Maniatis, T. (1983). Identification of two distinct regulatory regions adjacent to the human B-interferon gene. Cell **34**, 223-232.

Chapter 3

The differential regulatory capacities of c-myc and its pairing partner max

Sean V. Tavtigian and Barbara J. Wold

Introduction

More than a decade of work with myc gene family members has uncovered important roles in normal cellular proliferation, differentiation, and oncogenesis (for reviews, see Bishop 1986; Cole 1986). But cell biological measures of myc regulatory function have not yet been connected with explanatory molecular level functions. Several lines of evidence support the long standing hypothesis that myc proteins are transcription factors: (1) Conditional expression of c-myc or N-myc (Armelin et al., 1984; Cavalieri and Goldfarb 1988; Eilers et al., 1991) in growth arrested cells stimulate progression to S-phase. (2) c-Myc is expressed throughout the 14 hour G_1 that follows mitogen stimulation of quiescent cells, which suggests that myc proteins play a role in G_1 before the transition to S-phase (Muller et al., 1984; Pardee 1989). This argues for a role that precedes direct involvement in DNA replication, the other postulated mechanism of myc function. (3) At the molecular level, myc proteins are members of a transcription factor family defined by adjacent basic, helix-loop-helix (Murre et al., 1989), and leucine repeat (Landschulz et al., 1988) motifs, responsible for DNA binding, oligomerization, and oligomerization with an emphasis on specificity (Hu et al., 1990; O'Shea et al., 1992), respectively. In vitro experiments have defined an E box DNA sequence (CACGTG) for which myc proteins have specific binding affinity (Blackwell et al., 1990; Kerkhoff et al., 1991; Prendergast and Ziff 1991). (4) Finally, the expression of several genes whose promoters contain copies of the myc target E box were found to be regulated by conditional myc expression (chapter 2).

Among the bHLH, bzip, and bHLHzip protein families are individual members that efficiently form DNA binding homodimers, e.g., GCN4 (Kouzarides and Ziff 1989; Sellers and Struhl 1989), and members that bind DNA either preferentially or exclusively as heterodimers, e.g., MyoD (Lassar et al., 1991). Myc family proteins are members of the

latter class. To date, one gene, called max or myn, specifies at least four distinct bHLHzip products (Blackwood and Eisenman 1990; Prendergast et al., 1991; Makela et al., 1992) all of which can form DNA binding heteromers with myc family members in in vitro assays. Also, myc and max proteins are associated in vivo to the extent that myc is quantitatively coprecipitated by anti max sera, but the converse is not true (Wenzel et al., 1991; Blackwood et al., 1992).

Max differs from myc in several significant features. In addition to forming heteromers with myc, max can efficiently form DNA binding homodimers. As a homodimer, max DNA binding may be regulated by phosphorylation, but DNA binding by myc:max heteromers appears not to be (Berberich and Cole 1992). Unlike myc proteins, max proteins are relatively stable (Blackwood et al., 1992). This stability implies that low mRNA levels can support the accumulation of significanct levels of max protein; in contrast, c-myc protein accumulates and decays in rapid response to changes in the abundance of its mRNA. Max transcript levels are modestly induced during mid-G1 of the first cell cycle following serum stimulated emergence from growth arrest, i.e., the gene encodes a mid serum response (mid-SR) class transcript (Prendergast et al., 1991; chapter 2), but both the transcript and protein are present at appreciable levels throughout the cell cycle (Blackwood et al., 1992). Finally, the max proteins are very small, and no data published to date suggest that they posess a transcription activation domain.

Previously, we identified a number of myc regulated transcripts among a panel of mid-SR class genes (chapter 2). Their identification as myc regulated transcripts was made by showing that they are either up or down-regulated in response to induction of conditional myc expression in G_0 arrested 3T3 cells. In that experimental setting, the only source of max activity was supplied by the endogenous max gene. The degree to which regulation by myc depended on the presence of max, or was modulated by shifts in the ratio of max to myc, was unknown. Here we investigate this possibility using cell lines that provide altered max

expression either constitutively or conditionally. In addition, we present cell lines that supplement max levels in previously described conditional myc expressing backgrounds (chapters 1 and 2; Eilers et al., 1991). These are used to test: (1) the cell cycle progression effects of varied levels of the two regulators, (2) their effects on the expression of previously identified myc regulated mid-SR class genes, and (3) their effects on the expression of several immediate early serum response class genes.

Materials and methods.

Max expression vector. The coding sequences for max and max-9 were cloned by amplification from first strand cDNA derived from serum stimulated Balb/c 3T3 cells. Synthetic oligonucleotide primers were:

max 5': GCG GCT AGC GAT AAC GAT GAC ATC GAG

max 3': GCG TCG ACT CAG CTG GCC TCC ATC CGG A[G/A][C/T] TT[C/T] TT

A synthetic Bam HI - Xba I fragment (here called LE for linker epitope) containing a concensus translation initiator (Kozak 1988) and encoding the influenza HA1 epitope recognized by the monoclonal antibody 12CA5 (Wilson et al., 1984) was ligated to the Nhe I - Sal I max and max-9 fragments. Subsequently, the Bam HI - Sal I fragments encoding LEmax and LEmax-9 were used to replace the Bgl II - Xho I c-myc encoding fragment of pmyc23c (see chapter 1), generating the metallothionein-I promoter driven conditional max expression vectors pLEmax and pLEmax-9.

Cell culture. Cell lines were maintained in phenol red free DME (Gibco) that had been adjusted to 4mM hepes with the addition of dry hepes free acid. Medium supplements were 5% calf serum (Gibco), 1x glutamine penn-strep (Irvine Scientific), and 1x GMS-X supplement (Gibco). G418 resistant cell lines were maintained in medium containing 100 µg/ml G418 (Gibco). Hygromycin resistant cell lines were maintained in medium supplemented with 100 µg/ml Hygromycin B (Boehringer Manheim). Cells were cultured exclusively on fibronectin coated plates; these were either purchased from Collaborative Research or prepared by coating dishes (6 cm, 10 cm, 15 cm and 24 well, Nunc) for eight hours to overnight at room temperature with bovine fibronectin (15 µg/ml in PBS, Sigma).

Selection of max expressing cell lines. NIH 3T3, myc23c expressing 3-5B (see chapter 1), Balb/c 3T3, and c-myc estrogen receptor fusion protein expressing MER #6 cells (Eilers et al., 1991) were seeded

on fibronectin coated 6 well plates at a density of $5x10^4$ cells per well. 24 hours later, the cells were rinsed twice with transfection medium (1:1 phenol red free DME:phenol red free F12, 1x glutamine, 0.1x penn-strep, 1x GMS-X) and then transfected with the addition of 1 ml of transfection medium containing 100 ng of drug resistance marker, 900 ng of max encoding plasmid, and 17 µg of TransfectAce cationic liposomes (BRL). NIH 3T3 and Balb/c 3T3 were transfected with the G418 resistance marker pNeo 3 (Bond and Wold 1987) plus either pLEmax or pLEmax-9. 3-5B and MER #6 cells, which are already G418 resistant, were transfected with the hygromycin resistance marker pY3 (Blocklinger and Diggelmann 1984) plus either pLEmax or pLEmax-9. After 5 hours, the medium was adjusted to 5% calf serum with the addition of culture medium containing 10% calf serum. Drug resistant cell lines were selected in either 270 µg/ml G418 or 200 µg/ml hygromycin plus 100 µg/ml G418. After resistant colonies were established as lines, they were maintained as indicated in cell culture.

Max expressing cell lines were selected by immunostaining with the anti-HA epitope mAb 12CA5. Candidate cell lines were seeded in adjacent wells of fibronectin coated 24 well plates, mock or zinc stimulated for 4 hours, PBS rinsed, and fixed in 4% formaldehyde, 1x PBS for 10 minutes. Cells were blocked for 20 minutes in 4% goat serum, 0.4% Triton-X100, 1% BSA, 1x PBS; incubated with 1^o Ab overnight at 4^o in 1% goat serum, 0.4% Triton-X100, 1% BSA, 1x PBS; incubated with 2^o (biotinylated goat anti-mouse, Vector Laboratories) Ab for 1 to 2 hours in 0.02% Triton-X100, 0.25% BSA in 1x PBS; incubated for one hour with streptavidin biotin peroxidase complex (Amersham) in 1x PBS; and stained with nickel/cobalt enhanced DAB.

Cell cycle progression experiments. Cells were seeded at a density of approximately $3x10^4$ cells/well in fibronectin coated 24 well plates. 24 hours later the cell were refed with 0.75 ml medium. 72 hours later the cells were zinc induced with the addition of various

concentrations of β-estradiol, zinc sulfate, or both in phosphate free MEM (Gibco), mock induced with the addition of an appropriate volume of phosphate free MEM, or serum induced with the addition of FBS to 15% final concentration. The cells typically formed confluent monolayers about 48 hours after seeding. In both procedures, entry into S-phase was monitored by adding 37.5 µl of a labelling mixture containing 5-bromodeoxyuridine and 5-fluorouracil (Amersham) 14 hours after stimulation. Experiments were terminated 10 hours after labelling mix addition by washing the cells with PBS and then fixing them in 5% acetic acid in ethanol. Immunoperoxidase staining employed the Amersham cell proliferation kit. Wells were washed 3 x 3 minutes with PBS, incubated with 8-fold diluted anti BrdU mAB / nuclease mix for 4 hours at room temperature, washed 3 x 3 minutes with PBS, incubated with 4-fold diluted peroxidase labelled anti mouse Ab for 2 hours at room temperature, washed 3 x 3 minutes with PBS, stained with nickel/cobalt enhanced DAB, and counterstained with fast green FCF.

RNA preparation. Cells were seeded at a density of approximately $2x10^6$ cells/plate on fibronectin coated 10 cm plates. 24 hours later, they were refed with 10 ml of medium; the cells were subsequently brought to growth arrest and zinc, estrogen, or mock stimulated as in the cell cycle progression experiments. cells were harvested in 4M guanidinium thiocyanate, 25 mM sodium citrate pH 7.0, 0.5% N-lauryl sarcosine, and 100 mM 2-mercaptoethanol. 0.1 volumes of 2M sodium acetate pH 4.0, 1 volume of phenol, and 0.2 volumes of chloroform were added sequentially with vigorous mixing following each addition. The preparations were chilled in an ice bath for 15 minutes and then centrifuged at 10,000 rpm for 10 minutes at 4° in a SS34 rotor. The aqueous phase was reextracted with 1 volume of 1:2 phenol:chloroform, precipitated twice, and redissolved in TE (Chomczynski and Sacchi 1987).

Transcript measurement. RNA probes were synthesized using T7, T3, or SP6 RNA polymerase essentially as recommended by the enzyme

manufacturer. Full length probes were excised from 5% acrylamide gels crosslinked with BAC (N,N'-bis-acrylylcystamine, Bio-Rad) and the acrylamide was dissolved with 200 mM 2-mercaptoethanol in TE. RNase protections were performed by a modification of the procedure described by Zinn et al. (Zinn et al., 1983). Two or five µg of total cellular RNA plus tRNA to make a total of 25 µg were coprecipitated with probe. Samples were redissolved in 30 µl of RNA hybridization buffer,(40 mM Pipes pH 6.7, 400 mM NaCl, 1 mM EDTA, 80% formamide) denatured at 80° for 10 minutes, and hybridized for >12 hours. Probe concentrations and hybridization temperatures were optimized so that hybridizations proceeded to at least 90% of completion in 12 hours. Hybridizations were digested with the addition of 300 µl of 1500 units/ml RNase T1 (BRL) in RNase digestion buffer (10 mM Tris pH 7.5, 300 mM NaCl, 5mM EDTA) for 60 minutes at 30°, followed by 200 µl of 250 µg/ml proteinase K, 1% SDS for 15 minutes at 30°. Handling of samples containing dissolved BAC acrylamide was facilitated by co-precipitating RNA samples and probe out of 100 mM 2-mercaptoethanol and using well silanized tubes. Samples were prepared for electrophoresis, fractionated on 5% denaturing acrylamide gels, dried, and exposed to X-ray film. Count per minute data were also obtained from dried gels on a phosphorimager (Molecular Dynamics) and converted to transcript per cell data using internal CPM standards.

Results

One goal of this work was to compare and contrast the regulatory consequences of conditional myc expression with the regulatory consequences of conditional max expression and the regulatory consequences of simultaneous expression of both proteins. It also seemed possible that the response of particular myc regulated genes identified previously might have been limited by max availability. By testing these genes in different myc and max environments, we hoped to identify target:regulatory condition combinations where the magnitude of response is especially robust and therefore suitable for subsequent biochemical level analyses.

In previous work we identified a number of transcripts regulated by myc (see chapter 2). The conditional myc expressing systems used to detect myc responsiveness were 3-5B cells, a NIH 3T3 based cell line that expresses a metallothionein-I promoter driven c-myc gene (myc23c, see chapter 1) and MER-#6 cells, a Balb/c 3T3 based cell line that expresses a c-myc estrogen receptor fusion protein (mycER, see Eilers et al., 1989; Eilers et al., 1991). In this study, we have expanded the test cell repertoire to include two of the four known forms of max (Blackwood and Eisenman 1990; Prendergast et al., 1991; Makela et al., 1992). These alternative max forms are thought to result from: (1) an alternative splicing event that inserts 9 amino acids near its N-terminus, yielding max-9 (Blackwood and Eisenman 1990; Prendergast et al., 1991), and (2) an independent alternative splice near the C-terminus of the leucine repeat, resulting in synthesis of c-terminal truncated proteins called Δmax and Δmax-9 (Makela et al., 1992). The latter two max species were discovered while the work reported here was in progress and were therefore not tested.

Preparation of max expressing cell lines. To facilitate selection of max expressing cell lines and subsequent biochemical studies, we

prepared metallothionein-I promoter driven conditional expression vectors for HA1 epitope tagged (Wilson et al., 1984) max and max-9, termed LEmax and LEmax-9, respectively. These were stably cotransfected with either G418 or hygromycin B resistance into NIH 3T3, 3-5B, Balb/c 3T3, and MER #6 cells. Transfections and cell culture were carried out exclusively on fibronectin coated plates. The resulting cell lines were maintained at relatively low density and were always refed within 24 hours of splitting until they were well established in culture. These handling precautions were suggested by preliminary studies of cell lines expressing ectopic myc and max polypeptides.

Cell lines expressing the tagged max proteins were identified by immunostaining mock and zinc induced cells for the presence of the HA1 epitope. The max expressing cell lines subsequently identified shared with previously isolated myc and max expressing cell lines (data not shown) the characteristics of growth arrest at relatively low density and a propensity for fibronectin coated plates. Names and genotypes of both conditional max and myc plus max co-expressing cell lines are presented in table 1.

Cell line characterization

Regulatory effects on cell cycle progression. Conditional expression of myc in growth arrested cells can lead to emergence from growth arrest and progression through S-phase (Armelin et al., 1984; Cavalieri and Goldfarb 1987; Eilers et al., 1991); this is true both in cells that have been growth arrested at low density in serum free defined medium and in cells that have been growth arrested by contact inhibition in medium containing 5% calf serum (see chapter 2, also Hevezi and Bishop, unpublished). We found that survival of the myc plus max co-expressing cell lines was very poor in the defined medium growth arrest system (Zhan and Goldfarb 1986, see also chapters 1 and 2) used in previous work. On the other hand, over the timecourse required for

Derived cell line	Proximal parental line	myc transgene	max transgene	Conditional progression: basal	induced
NIH 3T3	NIH 3T3			0	0
30-1	NIH 3T3		LEmax	0	0
31-4	NIH 3T3		LEmax-9	5	5
3-5B	NIH 3T3	myc23c		30	75
37-3	3-5B	myc23c	LEmax	20	(both) 50
38-8	3-5B	myc23c	LEmax-9	20	(both) 65
Balb/c 3T3	Balb/c 3T3			0	0
32-7	Balb/c 3T3		LEmax	0	0
34-3	Balb/c 3T3		LEmax-9	5	5
MER #6	Balb/c 3T3	myc ER		5	55
36-7	MER #6	myc ER	LEmax-9	5	(max) 10 (myc) 80 (both) 80

Table 1. Summary of cell lines presented in this work. The first four columns give the name and genotype of each cell line used in this work. For the progression experiments presented in the last two columns, cells were seeded and brought to growth arrest in medium containing 5% calf serum as indicated in materials and methods. Arrested cells were mock, zinc (20 μM), β-estradiol (100 nM), or zinc plus β-estradiol stimulated as appropriate to the conditional myc and/or max genes expressed by each line. Cells were BrdU labelled from 14 to 24 hours following stimulation. Fixation and immunoperoxidase staining were as in materials and methods. Approximately 200 cells were counted for each condition; data are rounded to the nearest 5%. The conditional progression (index) is the number of BrdU stained cells divided by the total number of cells.

our experiments, the co-expressing cell lines tolerated well contact inhibition in medium containing 5% calf serum. Therefore, we chose this culture arrest condition to compare both the cell cycle progression effects and mRNA regulatory capacities of conditional max expression, conditional myc expression, and max-supplemented conditional myc expression.

The results of cell cycle progression experiments are presented in columns 5 and 6 of table 1. Several results stand out: (1) As measured by the fraction of cells incorporating BrdU, MER #6 and derived cell lines were subject to efficient growth arrest by contact inhibition in medium containing 5% calf serum. However, while 3-5B cells were growth arrested efficiently at low density in serum free defined medium (chapter1, chapter2), 3-5B and its max expressing derivatives were much less efficiently arrested by contact inhibition in medium containing 5% calf serum (compare an arrested labelling index of ~2% in defined medium to 20-30% in serum containing medium). (2) Transgene max expression alone did not stimulate progression to S-phase in any of the cell lines tested. (3) The Balb/c 3T3 based mycER plus LEmax-9 expressing cell line 36-7 is subject to efficient growth arrest by contact inhibition and reenters the cell cycle with unprecedented efficiency among myc lines (labelling index of 80%) following activation of conditional myc. In comparison with the parental MER #6 cells, this represents a significantly increased progression efficiency (80% versus 55%) and approaches the typical response (95%) to full serum stimulation. Since the difference between 36-7 and MER #6 cells is detectable with or without up-regulation of transgene max expression, the difference could be either clonal variation or a consequence of basal max overexpression from the transgene. (4) Basal expression of max alone from the Mt-I promoted transgene appears to promote and stabilize growth arrest. The observation is that monolayers of myc and myc plus max co-expressing cell lines pass through a stage in which they are relatively tightly packed, continually shed floating cells, and eventually become sub-

confluent. In contrast, max expressing cell lines form monolayers at lower density than parental NIH and Balb/c 3T3 cells. Over long periods of arrest, these monolayers appear to be more stable than monolayers of parental cells.

Auto and cross regulation involving myc and max. The longstanding observation of myc autoregulation (Adams et al., 1983; Nishikura et al., 1983) suggested that we should also look for max autoregulation and cross-regulation between myc and max. Our stably transfected cell lines should allow us to use the endogenous myc and max genes as reporters of these potential phenomena. Table 2 presents a survey of endogenous c-myc and total max transcript levels across the set of cell lines; the hsp 70 transcript MHS 214 (Lowe and Moran 1986) serves as a sensitive control for any stress response to the induction conditions. In this experiment, the myc transgene and endogenous RNAs were readily discriminated. However, the max riboprobe that we used did not efficiently distinguish between natural and transgene max; consequently, max levels are the sum of the transgene and endogenous RNA species.

As expected from prior immunocytochemical screening of the max cell lines, all expressed the transgene max RNA. Significant conditional induction above basal levels was achieved in the LEmax-9 cell lines 31-4 (5-fold), and 34-3 (12-fold). In contrast, the LEmax lines (30-1, and 32-7) expresses no more than 2-fold more total max RNA than does their parental line.

A striking and unexpected observation was that max up-regulation powerfully induces c-myc expression. In three of the max transgene expressing lines (30-1, 31-4, and 32-7) the natural c-myc transcript was overexpressed 2- to 3-fold under both basal and induced conditions. In the fourth conditional max cell line, c-myc RNA initiated from the dominant P2 promoter was induced 20-fold in response to conditional max expression. This phenomenon is clearly important for the interpretation

Table 2. Expression of myc, max, and hsp 70 in conditional myc and/or max expressing cell lines. Cells were seeded and brought to growth arrest in medium containing 5% calf serum as indicated in materials and methods. Cells were harvested for RNA preparation 8 hours following induction. For Balb/c 3T3 cells and derivitives, induction conditions were 100 nM β-estradiol or 20 μM zinc. For NIH 3T3 cells and derivitives, the induction condition was 25 μM zinc. RNAs were analyzed by ribonuclease protection. Quantitative data from each set of protections were obtained from the dried gels on a phosphorimager (Molecular Dynamics). Transcript prevalences were calculated from internal CPM standards on each gel.

Terminology:

1. Transcript abundance data are given in transcripts per cell, calculated as:

$$\frac{\text{protected cpm } (\mu\text{g total RNA}^{-1})}{\text{protected probe activity, cpm mol}^{-1}} \times \frac{1\ \mu\text{g total RNA}}{50{,}000\ \text{cells}} \times 6.02 \times 10^{23}\ \text{molecules mole}^{-1}$$

corrected for probe decay over the course of the experiment. More accurately, this is a measure of transcripts per 20 pg of total RNA; in our experience, a reasonable estimate of the RNA content of growth arrested 3T3 cells.

2. Normalized fold change is:

$$\frac{\text{fold change upon induction in conditional cell line}}{\text{fold change upon induction in control 3T3 cell line}}$$

3. Ratio to parental is:

$$\frac{\text{transcript prevalence in induced conditional cell line}}{\text{transcript prevalence in induced 3T3 control cell line}}$$

Transcript	Balb/c 3T3 based lines					NIH 3T3 based lines					
cell line:	Balb/c	32-7	34-3	MER	36-7	NIH	30-1	31-4	3-5B	37-3	38-8
max transgene:		LEmax	LEmax-9		LEmax-9		LEmax	LEmax-9		LEmax	LEmax-9
myc transgene:				mycER	mycER				myc23c	myc23c	myc23c
myc23c											
basal	not expressed in these cell lines								46	87	74
zinc treated									270	430	310
fold Δ									6.0	5.0	4.2
max											
basal	31	39	27	33	160	8.4	17	61	15	21	72
zinc treated	42	48	420			9.2	17	330	23	49	68
estrogen treated	40			44	120						
normalized Δ	(1.0)	0.91	12	1.0	0.57	(1.0)	0.95	4.9	1.4	2.1	0.85
ratio to parental		1.1	10	1.1	2.9		1.9	35	2.5	5.3	7.3
c-myc											
basal	1.7	4.9	1.5	2.7	3.5	2.4	5.7	4.8	4.5	2.6	3.2
zinc treated	2.3	6.0	41			3.2	6.6	7.6	1.9	2.1	1.9
estrogen treated	2.1			2.0	2.5						
normalized Δ	(1.0)	0.91	21	0.60	0.58	(1.0)	0.87	1.2	0.32	0.61	0.45
ratio to parental		2.6	18	1.0	1.2		2.1	2.4	0.59	0.66	0.59
hsp 70											
basal	7.2	9.4	8.5	7.3	11	3.1	2.9	3.8	3.7	4.4	6.1
zinc treated	8.1	8.4	11			4.2	2.9	6.2	2.6	4.5	6.1
estrogen treated	7.3			5.2	8.2						
normalized Δ	(1.0)	0.79	1.1	0.70	0.76	(1.0)	0.74	1.2	0.52	0.75	0.74
ratio to parental		1.0	1.4	0.71	1.1		0.69	1.5	0.62	1.1	1.5

of subsequent effects, as high max levels are now appreciated to carry with them elevated c-myc.

Endogenous c-myc transcript levels have previously been shown to be subject to transcriptional autorepression (Cleveland et al., 1988; Grignani et al., 1990; Penn et al., 1990). Here, that phenomenon was reflected by 1.5- to 3-fold repression of the natural c-myc transcript in response to conditional myc activation in each of the relevant lines.

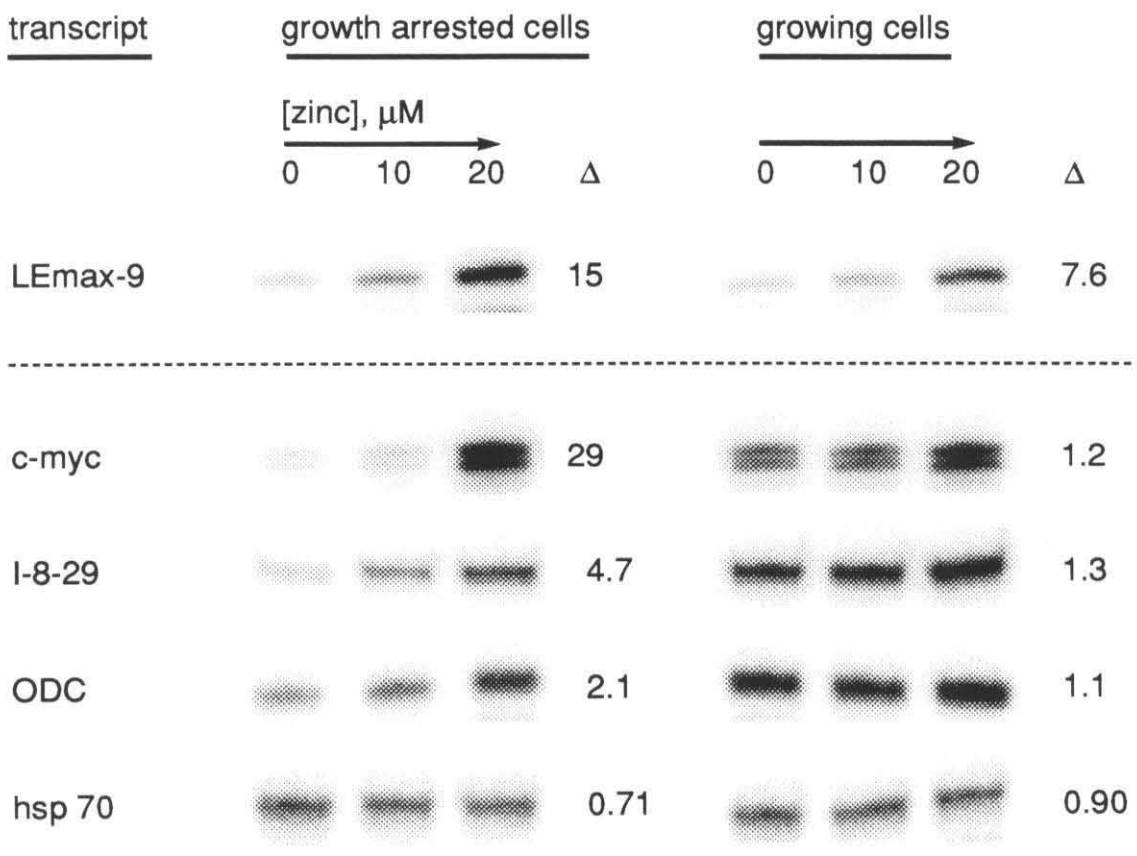
Finally, the hsp 70 control was neither basally overexpressed nor significantly induced under any of the regulatory conditions presented (table 2).

Regulation by max expression.

The observation that up-regulation of the c-myc transcript occured in all of the max overexpressing cell lines motivated further investigation of max as a possible regulator of other immediate early and mid-SR class genes. Using the LEmax-9 expressing Balb/c 3T3 line 34-3, we first asked how regulation by conditional max expression depended on the growth state of the cells. In figure 1A we compare the effects of conditional max expression in arrested and growing cells, using one immediate early serum response transcript, c-myc, and two mid-SR transcripts, I-8-29 and ornithine decarboxylase, as reporters. The hsp 70 transcript MHS 214 served as control. The result was clear; all three serum regulated transcripts responded strongly to conditional max expression in growth arrested cells but showed little to no response in exponentially growing cells. Furthermore, the difference between responses in arrested and growing cells was not due to performance of the conditional expression system; LEmax-9 was sharply induced in both cases.

Figure 1B presents normalized uninduced transcript levels for the induced reporters c-myc, ornithine decarboxylase, and I-8-29; tenascin, which was overexpressed but not induced; and hsp 70. Here there were

Figure 1


transcript
growth arrested cells
growing cells
[zinc], µM
0
10
20
Δ
0
10
20
Δ
LEmax-9
15
7.6
c-myc
29
1.2
I-8-29
4.7
1.3
ODC
2.1
1.1
hsp 70
0.71
0.90

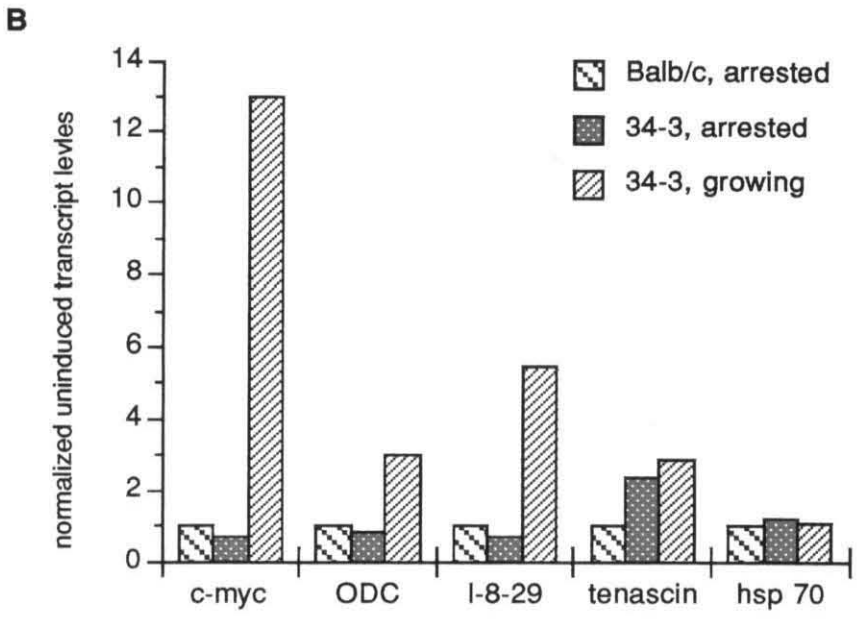


Figure 1. Comparison of conditional max expression in G_0 arrested and growing cells. **A.** Transcript reponses to conditional max expression. Zinc induction: 0 = mock, 10 = 10 μM zinc, 20 = 20 μM zinc, Δ = normalized fold induction at 20 μM zinc (see definition in table 2 legend). **B.** Comparison of basal transcript levels in G_0 arrested and growing cells. Data are normalized to transcript levels in G_0 arrested Balb/c 3T3 cells.

Culture conditions. G_0 arrested cells: Balb/c 3T3 cells were seeded in medium containing 5% calf serum at a density of 1.0×10^6 cells per plate. 34-3 cells were seeded in medium containing 5% calf serum at a density of 7.5×10^5 cells per plate. The cells were refed 24 hours later and zinc stimulated 72 hours after that. Growing cells: 34-3 cells were seeded in the same medium at a density of 1.0×10^5 cells per plate. They were refed 48 hours later and zinc stimulated 48 hours after that.

RNA preparation. In all cases cells were harvested 8 hours after induction. mRNAs were analyzed by ribonuclease protection. Quantitative data from each set of protections were obtained from the dried gels on a phosphorimager (Molecular Dynamics).

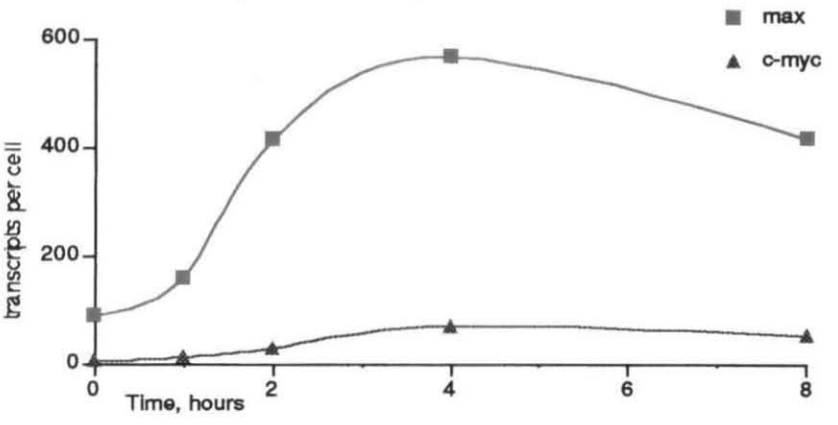
two correlations that provide some explanation of the differential response to conditional max expression between arrested and growing cells. First, those transcripts that were expressed at higher levels in growing than arrested 34-3 cells were also inducible by conditional max expression. Second, for the inducible reporters, transcript levels in growth arrested 34-3 cells were at or below transcript levels in arrested Balb/c cells. However, tenascin in arrested 34-3 cells, and all of the reporter transcript levels in growing 34-3 cells, were higher than in arrested Balb/c cells. Thus basal overexpression, whether due to growth state or a difference between the cell lines, correlated with reduced or abrogated responsiveness to conditional max expression

Because the immediate early serum response gene c-myc responded more strongly to conditional max expression than did any of the mid-SR transcripts examined, we examined the timecourse of response of other immediate early serum response transcripts to the regulatory pattern initiated by conditional max expression in growth arrested 34-3 cells (figure 2). Panel A presents total transcript levels for max and c-myc over the timecourse of the experiment; of three RNA level experiments performed with this cell line, this one showed the smallest regulatory responses to max expression that we observed. Even so, c-myc transcript levels were induced 11-fold by 4 hours after stimulation of max expression. Furthermore, no response was detected from either the hsp 70 transcript MHS 214 or the c-fos transcript, both of which are extremely sensitive to the toxic effects of heavy metal poisoning (c-fos is shown in panel B, hsp 70 data are not shown). This is strong evidence that the observed responses were to the myc and max regulators, not a zinc stress response. In addition, the expected p23 LEmax-9 protein synthesized during the 2 to 4 hour induction window is detected by immunoprecipitation (appendix A).

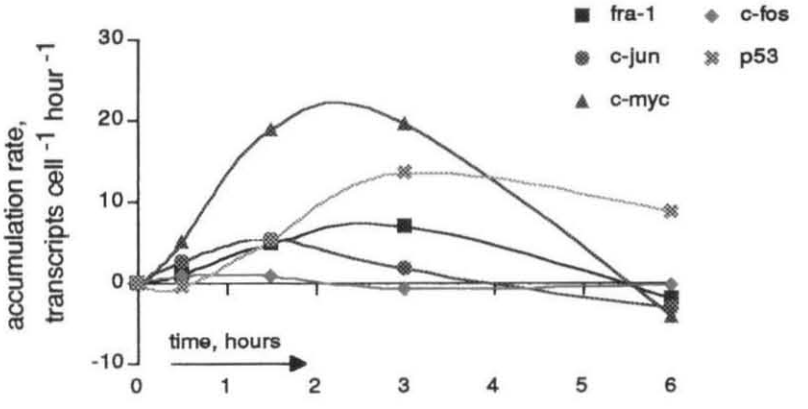
Panel B presents transcript accumulation rates for the transcriptional regulators c-fos, fra-1, c-jun, c-myc, and p53 tumor antigen. Like c-myc, the immediate early serum response class AP-1

Figure 2. Response of a subset of immediate early and mid-SR class transcription factor transcripts to conditional max expression in growth arrested cells. 34-3 cells were seeded and brought to growth arrest in medium containing 5% calf serum as indicated in materials and methods. Arrested cells were mock or zinc (20 μM) induced over an 8 hour timecourse and then harvested for RNA preparation. RNAs were analyzed by ribonuclease protection. Quantitative data from each set of protections were obtained from the dried gels on a phosphorimager (Molecular Dynamics). Transcript prevalences were calculated from internal CPM standards on each gel. **A.** Total transcript prevalence for max and c-myc. **B.** Transcript accumulation rates (calculated as [Δ transcripts $cell^{-1}$ / Δ time] over each interval) for the transcription factor transcripts c-myc, c-fos, c-jun, fra-1, and p53 tumor antigen. **C.** Transcript accumulation rates for the mid-SR transcripts G1 cyclin D2, G1 cyclin D3, I-8-29, and p53 tumor antigen.

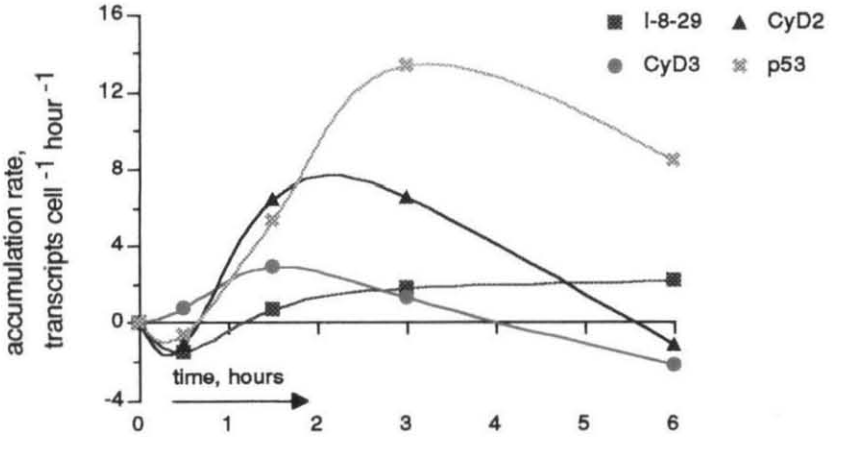
A. Total transcipt levels for myc and max



B. Transcipt accumulation rates for selected transcription factors



C. Transcipt accumulation rates for selected mid-SR transcripts



family members fra-1 and c-jun also responded, but c-fos did not. The kinetics of c-jun expression were quite fast; its accumulation rate peaked within 90 minutes of the point at which max expression was stimulated. Fra-1 and c-myc responded with slightly slower kinetics, and the response of the p53 tumor antigen transcript was reminiscent of mid-SR kinetics. Panel C presents transcript accumulation rates for the mid-SR genes p53 tumor antigen, I-8-29, G_1 cyclin D2, and G_1 cyclin D3. While the timecourses of the p53 tumor antigen and I-8-29 transcript responses resembled those expected for mid-SR genes, the G_1 cyclin timecourses do not. In particular, the G_1 cyclin transcripts responded early and then decayed away as the cells approached (in a temporal, not a biochemical sense) what would be the restriction point if they had been driven into a productive cell cycle.

Regulation of other transcripts by conditional myc expression with or without augmented max.

The temporal pattern of c-myc and max expression during the first cell cycle following serum stimulated emergence from growth arrest suggests that the activity of a potential myc:max heteromeric transcription factor will peak during the extended G_1 of that cycle (see chapter 2, also Pardee 1989; Prendergast et al., 1991; Blackwood et al., 1992). To identify transcripts that are regulatory targets of c-myc, we previously cloned a panel of 18 mid-SR clones. From that group of clones we identified a subset that are subject to regulation by c-myc (chapter 2). In this work we pursued the regulation of a subset of those clones consisting of the anonymous cDNA I-8-29, lactate dehydrogenase, heat stable antigen, ornithine decarboxylase, and tenascin. In a conceptually similar cloning exercise, the G_1 cyclins D1, D2, and D3 were identified as members of the mid-SR class (Matsushime et al., 1991); we subsequently found that G_1 cyclin D2 is

regulated by c-myc. Finally we followed the expression of another myc regulated gene, prothymosin α (Eilers et al., 1991; chapter 2).

In table 3 we compare the regulation of this panel of mRNAs in cell lines that conditionally express myc (MER #6 and 3-5B) with derivative cell lines that, in addition to conditionally expressing a myc protein, basally overexpress LEmax-9 (36-7 and 38-8). The main issues in surveying the effects of conditional myc expression on the members of this set of target genes are: (1) The magnitude of response of individual targets to induced myc expression under this arrest condition. (2) Possible differences between that response in this arrest condition and the low density defined medium arrest condition of chapter 2. (3) The consequences of added max expression upon either absolute reporter transcript levels or response to conditional myc expression.

Up-regulation in response to myc expression. The magnitude of the response of myc up-regulated genes in the Balb/c^MER #6 system was consistently equal to or greater than it was when those responses were measured in the culture condition of growth arrest at low density in defined medium (see chapter 2). On the other hand, the magnitude of target gene responses to conditional myc expression in the NIH^3-5B system was quite similar in both culture conditions. Supplementation of max expression in the conditional myc expressing cell lines had substantial effects on target expression, but those effects varied from target to target.

The regulatory pattern of I-8-29 provides an especially interesting example. In the Balb/c^MER #6^36-7 series, the transcript was 5.5- to 7.5-fold induced by conditional myc expression; in addition, the transcript was both more strongly induced and expressed at higher basal levels in LEmax-9 expressing 36-7 cells than in MER #6 cells. This is a striking case sensitivity to both myc and max; estrogen treated 36-7 overexpressed the I-8-29 transcript 62-fold relative to

estrogen treated Balb/c 3T3 cells. In the NIH^3-5B^38-8 series, the I-8-29 transcript was 5- to 6-fold basally overexpressed in both myc expressing cell lines, but its additional response to conditional myc expression was less than 2-fold.

G_1 cyclin D2 and lactate dehydrogenase followed a different regulatory pattern. G_1 cyclin D2 was induced 2- to 3- fold from basal levels in both MER #6 and 36-7 cells, i.e., there was no apparent effect due to added max expression. In the NIH system, this same transcript was induced just less than 2-fold in both 3-5B and 38-8 cells, but the basal level of its expression increased with increasing basal max expression. A similar pattern of regulation held for lactate dehydrogenase, except that it was slightly more strongly dependent on conditional myc expression in the Balb/c^MER #6^36-7 series and slightly less dependent on conditional myc expression in the NIH*3-5B*38-8 series.

Down-regulation in response to myc expression. Heat stable antigen is expressed at modest levels in Balb/c 3T3 cells, but is not expressed in the NIH 3T3 isolate in which 3-5B and 38-8 cells were built (other NIH 3T3 cells do express this transcript). Basal transcript levels for this gene were similar across the Balb/c^MER #6^36-7 series. The transcript was 2- to 3-fold repressed by conditional myc expression, and the magnitude of repression increased with added max. The regulatory response of the transcript for tenascin may be the most unexpected of all. The transcript is repressed by myc. In the Balb/c^MER #6^36-7 series this was reflected by basal repression in MER #6 cells relative to Balb/c 3T3 cells and by repression with conditional myc expression in 36-7 cells. In the NIH^3-5B^38-8 series, repression of the tenascin transcript by myc was observed only upon induction of myc expression. However, basal expression of the tenascin transcript was dramatically increased with the presence of the LEmax-9 transcript; basal expression was 7-fold higher in 36-7 cells than in

MER #6 cells, and basal levels increased with max expression across the NIH^3-5B^38-8 series so that they were 25-fold higher in 38-8 cells than in NIH 3T3 cells.

Transcript	Balb/c 3T3 based lines			NIH 3T3 based lines		
cell line:	Balb/c	MER	36-7	NIH	3-5B	38-8
max transgene:			LEmax-9			LEmax-9
myc transgene:		mycER	mycER		myc23c	myc23c
G1 cyclin D2						
basal	69	84	65	9.7	13	33
zinc treated				10	25	60
estrogen treated	71	190	180			
normalized Δ	(1.0)	2.2	2.7	(1.0)	1.8	1.7
ratio to parental		2.7	2.6		2.5	5.9
G1 cyclin D3						
basal	48	38	47	27	35	46
zinc treated				29	39	61
estrogen treated	48	51	68			
normalized Δ	(1.0)	1.3	1.5	(1.0)	1.0	1.2
ratio to parental		1.1	1.4		1.3	2.1
I-8-29						
basal	1.5	6.9	12	1.3	8.1	7.3
zinc treated	1.7			1.7[A]	14	16
estrogen treated	1.4	37	87			
normalized Δ	(1.0)	5.7	7.5	(1.0)	1.3	1.7
ratio to parental		26	62		8.3	9.2
lactate dehydrogenase						
basal	44	66	32	76	160	220
zinc treated	60			97[A]	390	370
estrogen treated	50	340	150			
normalized Δ	(1.0)	4.5	4.0	(1.0)	1.9	1.3
ratio to parental		6.7	3.0		4.0	3.8
ornithine decarboxylase						
basal	14	24	22	7.8	16	30
zinc treated				9.3	46	63
estrogen treated	14	81	62			
normalized Δ	(1.0)	3.3	2.8	(1.0)	2.5	1.8
ratio to parental		5.9	4.5		5.0	6.7

Transcript	Balb/c 3T3 based lines			NIH 3T3 based lines		
cell line:	Balb/c	MER	36-7	NIH	3-5B	38-8
max transgene:			LEmax-9			LEmax-9
myc transgene:		mycER	mycER		myc23c	myc23c
prothymosin α						
basal	860	980	620	270	880	660
zinc treated				340	1,300	1,200
estrogen treated	1,100	2,100	1,300			
normalized Δ	(1.0)	1.7	1.6	(1.0)	1.2	1.5
ratio to parental		1.9	1.2		3.8	3.6
tenascin						
basal	8.6	2.6	19	4.3	27	110
zinc treated				5.9	20	78
estrogen treated	8.4	2.2	12			
normalized Δ	(1.0)	0.85	0.66	(1.0)	0.52	0.53
ratio parental		0.26	1.5		3.3	13.1
heat stable antigen						
basal	34	38	37	this transcript is not		
zinc treated				expressed by NIH 3T3 cells		
estrogen treated	37	23	13			
normalized Δ	(1.0)		0.54	0.32		
ratio to parental		0.61	0.35			

Table 3. Expression of myc target genes in conditional myc expressing cell lines. Cells were seeded and brought to growth arrest in medium containing 5% calf serum as indicated in materials and methods. Cells were harvested for RNA preparation 8 hours following induction. For Balb/c 3T3 cells and derivitives, induction conditions were 100 nM β-estradiol or 20 μM zinc. For NIH 3T3 cells and derivitives, the induction condition was 15 μM zinc for the transcripts I-8-29 and lactate dehydrogenase (marked A) and 25 μM zinc for all of the others. Quantitation and analysis are as in table 2.

Discussion.

The regulatory effects of myc and max expression can be measured indirectly via a number of assays. They vary from inhibition of differentiation by myc to in vitro DNA binding by both max homodimers and myc:max heteromers at the E box sequence CACGTG. However, without knowing the direct effects of myc and max expression on their physiologic targets together with some understanding of those targets' function, it is not possible to connect myc and max biochemical functions with the biological effects of their expression.

Cell cycle progression and transformation. Provocative parallels can be drawn between: (1) Cotransformation by myc and activated ras proteins. (2) The consequences of adding max genes to the myc/activated ras cotransformation assay. (3) Cell cycle progression in response to conditional myc expression. (4) The effects of adding max to cell cycle progression experiments, and (5) The growth characteristics of max overproducing cell lines. Thus myc, originally recognized as an immortalizing oncogene, cooperates with activated transforming oncogenes, e.g., Ha-ras, in transformation assays such as rat embryo fibroblast focus formation assays (for reviews, see Bishop 1986; Cole 1986). At low concentrations, added max aids the activity of myc in the cotransformation assays (Prendergast et al., 1991); at higher concentrations max can become inhibitory, but Δmax maintains its cooperativity (Makela et al., 1992). (3), (4), and (5): Conditional expression of myc in growth arrested immortalized cell lines stimulates emergence from growth arrest and progression through S-phase (Armelin et al., 1984; Cavalieri and Goldfarb 1987; Eilers et al., 1991). Addition of transgene max to the conditional myc expressing cell line MER #6 can increase the efficiency of conditional cell cycle progression to near quantitative levels, but overexpression of max by itself is qualitatively anti progressive. Thus the effects of myc and max may be

parallel in the two assay systems, suggesting that they are both measures of the same underlying regulatory phenomenon.

Regulation of mRNA expression by myc and max. The expression timecourse of c-myc following serum stimulation of growth arrested cells, its nuclear localization, its DNA binding capacity (as a myc:max heteromer), and the presence of transcription factor specific structural motifs within its primary sequence (for review, see Prendergast and Ziff 1992) all point to a transcription factor function that should be most evident during mid G_1 of the extended first cell cycle following emergence from growth arrest. In support of this hypothesis, we found that about 1/3 of an assortment of mid serum response transcripts are in fact regulatory targets, direct or indirect, of c-myc (see chapter 2). The promoters of two of these genes, lactate dehydrogenase and ornithine decarboxylase, as well as the promoter of the myc regulated gene p53 tumor antigen, contain the E box to which myc:max heteromers and max homodimers can bind in vitro. Indeed, both myc:max heteromers and max homodimers will bind to an E box containing oligonucleotide derived from the lactate dehydrogenase promoter (appendix A). However, the presence within the cell of a number of transcription factors, including USF/MLTF and TFII-i (Carthew et al., 1985; Gregor et al., 1990; Roy et al., 1991) that can bind this same sequence complicates the implied conclusion that they are directly regulated targets.

Two related patterns become evident in summing across: (1) Messenger RNA up-regulation by conditional myc expression in cells that have been growth arrested at low density in defined medium (chapter 2). (2) Messenger RNA up-regulation by conditional myc expression or conditional myc with supplemented max expression in cells that have been growth arrested by contact inhibition in medium containing 5% calf serum. and (3) Messenger RNA up-regulation by conditional max expression in cells that are either proliferating in or have been growth arrested by contact inhibition in medium containing 5% calf serum. The first

emergent pattern is that when culture conditions are such that cell cycle reentry of growth arrested cells remains tightly dependent on conditional myc expression, the magnitude of mRNA responses to conditional myc expression are maintained across culture conditions (this follows from comparison of the Balb/c^MER 6 system in chapter 2, table 1A and the Balb/c^MER 6^36-7 system in table 3; a specific transcripts to examine is I-8-29). However, when the interaction between culture conditions and basal expression of the conditional myc transgene is such that the cells do not growth arrest efficiently, then the magnitude of myc dependent conditional mRNA regulation decreases (this follows from comparison of the NIH^3-5B system in chapter 2, table 1A and the NIH^3-5B^38-8 system in table 3; specific transcripts to examine ar prothymosin α and ornithine decarboxylase). The second pattern is that when the transcript levels of direct and indirect targets of max are similarly low in growth arrested Balb/c 3T3 and LEmax-9 expressing 34-3 cells, those transcripts subject to up-regulation in response to induced expression of max (c-myc, I-8-29, ornithine decarboxylase) respond with reasonable magnitude. In contrast, when the uninduced transcript level is higher in arrested 34-3 cells than in arrested Balb/c 3T3 cells (tenascin), or when transcript levels are elevated because the cells are proliferating, then the response to conditional max expression is decreased or eliminated. Thus in both primarily myc conditional and primarily max conditional settings, good growth arrest correlates with good target transcript up-regulation in response to conditional expression of the regulator. However, the degree to which this difference results from differential postranslational modification of the conditional regulator, versus functional redundancy in already growing cells, remains unknown.

The ability of high levels of max activity to interfere with myc ras cotransformation might superficially suggest that max would repress expression of myc activated genes. However, this was a pervasive pattern in neither the Balb/c^MER #6^36-7 nor the NIH^3-5B^38-8 cell

series. Rather, max up-regulated expression of a specific subset of the immediate early serum response transcript class. The most dramatic effect was on c-myc, but fra-1 and c-jun were also up-regulated while c-fos, the most strongly serum growth factor regulated member of that class (Muller et al., 1984; Lau and Nathans 1987; Almendral et al., 1988), was not. In addition two myc repressed genes, c-myc itself and tenascin, are induced by max overexpression. Unlike c-myc, the max protein is quite stable, not profoundly growth regulated, and present at appreciable levels even in growth arrested cells (Blackwood et al., 1992). If its functionality either alone, in a heteromeric complex with other non myc proteins, or in cooperation with other regulators of the immediate early serum response, is rapidly stimulated via the protein synthesis independent regulatory pathways activated by polypeptide growth factors, then it would in fact fit as a regulator of immediate early expression.

Future prospects. By definition, each of the myc and/or max regulated genes we have examined possesses it own unique promoter, splicing patterns, characteristics of mRNA stability, etc. Consequently, each of them responds to conditional myc and/or max expression with its own directional, magnitudinal, and temporal pattern. One goal of this work has been to identify those targets whose regulation by myc and/or max suggest a particularly promising entree to biochemical level analyses. Clearly, the most strongly myc up-regulated gene we have identified is I-8-29; a full cDNA sequence and analysis of the genomic clone are high priorities. Tenascin and c-myc both provide examples of transcripts that are repressed by conditional myc expression but up-regulated by max. c-myc is particularly amenable to analysis because we have both cDNA and genomic clones. Fra-1 provides a good example of a gene up-regulated by max but largely unregulated by myc alone. Both cDNA and genomic clones are in hand and can now be used to extend our analysis of its max response.

Studying transcript regulation by c-myc in a conditional environment where myc expression stimulates emergence from growth arrest and cell cycle progression poses a familiar problem in dissecting regulatory cascades: Are the targets we can identify directly regulated by myc, or are they indirectly regulated via the myc induced cell cycle? That molecular level question will be answered by direct molecular and biochemical studies of the individual regulators. However, the question itself implies that the pattern of up and down regulation controlled by myc is sufficient to emergence from G_0 and progression across G_1 to S-phase. Thus understanding the physiologic role of the myc proteins at a cell biological level will require continued cloning and functional analysis of these targets. Our indirect cloning approach (chapter 2) was successful but tedious. However, the difference in the levels of expression of I-8-29 in growth arrested Balb/c 3T3 cells and estrogen stimulated {mycER plus LEmax-9} 36-7 cells suggests that a direct subtractive cloning approach is now within reach.

Transcript up-regulation by conditional max expression poses the opposite problem: Why does max up-regulate expression of a variety of immediate early regulators but not stimulate cell cycle progression? Since max activates a distinct subset of immediate early class transcripts, for instancec-myc, fra-1, and c-jun, but not c-fos, the answer must lie in the pattern and downstream functions of those genes that are regulated. Happily, extensive libraries of immediate early probes are available (Lau and Nathans 1987; Mohn et al., 1991), rendering the first step of that problem immediately accessible to experimental analysis.

References

Adams, J. M., Gerondakis, S., Webb, E., Corcoran, L. M., and Cory, S. (1983). Cellular myc oncogene is altered by chromosomal translocation to an immunoglobulin locus in murine plasmacytomas and is rearranged similarly in human burkitt lymphomas. Proceedings of the National Academy of Sciences USA **80**, 1982-1986.

Almendral, J. M., Sommer, D., MacDonald-Bravo, H., Burckhardt, J., Perera, J., and Bravo, R. (1988). Complexity of the early genetic response to growth factors in mouse fibroblasts. Molecular and Cellular Biology **8**, 2140-2148.

Armelin, H. A. M., Armelin, C. S., Kelly, K., Stewart, T., Leder, P., Cochran, B. H., and Stiles, C. D. (1984). Functional role for c-myc in mitogenic response to platelet-derived growth factor. Nature **310**, 655-660.

Berberich, S. J., and Cole, M. D. (1992). Casein kinase-II inhibits the DNA-binding activity of max homodimers but not myc max heterodimers. Genes and Development **6**(2), 166-176.

Bishop, J. M. (1986). Viral Oncogenes. Cell **42**, 23-38.

Blackwell, T. K., Kretzner, L., Blackwood, E. M., Eisenman, R. N., and Weintraub, H. (1990). Sequence-specific DNA binding by the c-myc protein. Science **250**, 1149-1151.

Blackwood, E. M., and Eisenman, R. N. (1990). Max: A helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with myc. Science **251**, 1211-1217.

Blackwood, E. M., Luscher, B., and Eisenman, R. N. (1992). Myc and max associate in vivo. Genes and Development **6**(1), 71-80.

Blocklinger, K., and Diggelmann, H. (1984). Hygromycin B phophotransferase as a selectable marker for DNA transfer experiments with higher eukaryotic cells. Molecular and Cellular Biology **4**, 2929-2931.

Bond, V. C., and Wold, B. (1987). Poly-l-ornithine-mediated transformation of mammalian cells. Molecular and Cellular Biology **7**, 2286-2293.

Carthew, R. W., Chodosh, L. A., and Sharp, P. A. (1985). An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. Cell **43**, 439-448.

Cavalieri, F., and Goldfarb, M. (1987). Growth-factor-deprived BALB/c 3T3 murine fibroblasts can enter the S phase after induction of c-myc gene expression. Molecular and cellular biology **7**, 3554-3560.

Cavalieri, F., and Goldfarb, M. (1988). N-myc proto-oncogene expression can induce DNA replication in balb/c 3T3 fibroblasts. Oncogene **2**, 289-291.

Chomczynski, P., and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162**, 156-159.

Cleveland, J. L., Huleihel, M., Bressler, P., Siebenlist, U., Akiyama, L., Eisenman, R. N., and Rapp, U. R. (1988). Negative regulation of c-myc transcription involves myc family proteins. Oncogene Research **3**, 357-375.

Cole, M. D. (1986). The myc oncogene: its role in transformation and differentiation. Annual Review of Genetics **20**, 361-384.

Eilers, M., Picard, D., Yamamoto, K. R., and Bishop, J. M. (1989). Chimaeras of Myc oncoprotein and steroid receptors cause hormone-depndent transfromation of cells. Nature **340**, 66-68.

Eilers, M., Schirm, S., and Bishop, J. M. (1991). The myc protein activates transcription of the alpha-prothymosin gene. European Molecular Biology Organization Journal **10**(1), 133-141.

Gregor, P. D., Sawadogo, M., and Roeder, R. G. (1990). The adenovirus major late transcription factor USF is a member of the helix-loop-helix group of regulatory proteins and binds DNA as a dimer. Genes and Development **4**, 1730-1740.

Grignani, F., Lombardi, L., Inghirami, G., Sternas, L., Cechova, K., and Dalla-Favera, R. (1990). Negative autoregulation of c-myc gene expression is inactivated in transformed cells. European Molecular Biology Organization Journal **9**(12), 3913-3922.

Hu, Y.-F., Luscher, B., Admon, A., Mermod, N., and Tjian, R. (1990). Transcription factor AP-4 contains multiple dimerization domains that regulate dimer specificity. Genes and Development **4**, 1741-1752.

Kerkhoff, E., Bister, K., and Klempnauer, K. H. (1991). Sequence-specific DNA binding by myc proteins. Proceedings of the National Academy of Sciences USA **88**, 4323-4327.

Kouzarides, T., and Ziff, E. (1989). Leucine zippers of fos, jun, and GCN4 dictate dimerization specificity and thereby control DNA binding. Nature **340**(6234), 568-571.

Kozak, M. (1988). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Research **15**, 8125-8148.

Landschulz, W. H., Johnson, P. F., and Mcknight, S. L. (1988). The leucine zipper - a hypothetical structure common to a new class of DNA-binding proteins. Science **240**(4860), 1759-1764.

Lassar, A. B., Davis, R. L., Wright, W. E., Kadesch, T., Murre, C., Voronova, A., Baltimore, D., and Weintraub, H. (1991). Functional activity of myogenic HLH proteins requires hetero-oligerization with E12/E47-like proteins in vivo. Cell **66**(2), 305-315.

Lau, L. F., and Nathans, D. (1987). Expression of a set of growth-related immediate early genes in Balb/c 3T3 cells: coordinate regulation with c-fos or c-myc. Proceedings of the National Academy of Sciences USA **84**, 1182-1186.

Lowe, D. G., and Moran, L. A. (1986). Molecular cloning and analysis of DNA complementary to three mouse M_r = 68,000 heat shock protein mRNAs. The Journal of Biological Chemistry **261**(5), 2102-2112.

Makela, T. P., Koskinen, P. J., Vastrik, I., and Alitalo, K. (1992). Alternative forms of max as enhancers or suppressors of myc-ras cotransformation. Science **256**, 373-376.

Matsushime, H., Roussel, M. F., Ashmun, R. A., and Sherr, C. J. (1991). Colony-stimulating factor 1 regulates novel cyclins during the G1 phase of the cell cycle. Cell **65**, 701-713.

Mohn, K. L., Laz, T. M., Hsu, J.-C., Melby, A. E., Bravo, R., and Taub, R. (1991). The immediate-early growth response in regenerating liver and insulin-stimulated H-35 cells: comparison with serum stimulated 3T3 cells and identification of 41 novel immediate-early genes. Molecular and Cellular Biology **11**(1), 381-390.

Muller, R., Bravo, R., Burckhardt, J., and Curran, T. (1984). Induction of c-fos gene and protein by growth factors precedes activation of c-myc. Nature **312**, 716-720.

Murre, C., McCaw, P. S., and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, myoD1, and myc proteins. Cell **56**, 777-783.

Nishikura, K., Abbas, R., Erikson, J., Watt, R., Rovena, G., and Croce, C. M. (1983). Differential expression of the normal and translocated human c-myc oncogenes in B cells. Proceedings of the National Academy of Sciences USA **80**, 4822-4826.

O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1992). Mechanism of specificity in the fos-jun oncoprotein heterodimer. Cell **68**(4), 699-708.

Pardee, A. B. (1989). G1 events and regulation of cell proliferation. Science **246**, 603-608.

Penn, L. J. Z., Brooks, M. W., Laufer, E. M., and Land, H. (1990). Negative autoregulation of c-myc transcription. European Molecular Biology Organization Journal **9**(4), 1113-1121.

Prendergast, G. C., Lawe, D., and Ziff, E. B. (1991). Association of Myn, the murine homolog of max, with c-myc stimulates methylation-sensitive DNA-binding and ras cotransformation. Cell **65**(3), 395-407.

Prendergast, G. C., and Ziff, E. B. (1991). Methylation-sensitive sequence-specific DNA binding by the c-myc basic region. Science **251**, 186-189.

Prendergast, G. C., and Ziff, E. B. (1992). A new bind for myc. Trends in Genetics **8**(3), 91-96.

Roy, A. L., Pognonec, M. M. P., and Roeder, R. G. (1991). Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF. Nature **354**, 245-248.

Sellers, J. W., and Struhl, K. (1989). Changing fos oncoprotein to a jun-independent DNA-binding protein with GCN4 dimerization specificity by swapping leucine zippers. Nature **341**(6237), 74-76.

Wenzel, A., Cziepluch, C., Hamann, U., Schurmann, J., and Schwab, M. (1991). The N-myc oncoprotein is associated in vivo with the phosphoprotein max(p20/22) in human neuroblastoma cells. European Molecular Biology Organization Journal **10**(12), 3703-3712.

Wilson, I. A., Niman, H. L., Hougten, R. A., Cherenson, A. R., Connolly, M. L., and Lerner, R. A. (1984). The structure of an antigenic determinant in a protein. Cell **37**, 767-778.

Zhan, X., and Goldfarb, M. (1986). Growth factor requirements of oncogene-transformed NIH 3T3 and BALB/c 3T3 cells cultured in defined media. Molecular and Cellular Biology **6**, 3541-3544.

Zinn, K., Dimaio, D., and Maniatis, T. (1983). Identification of two distinct regulatory regions adjacent to the human B-interferon gene. Cell **34**, 223-232.

Conclusion

At the cell biological level, the functions of the myc family proteins will eventually be understood in terms of the identities and roles of the genes that they regulate and the various physiological contexts in which that regulation takes place. At the molecular level, the functions of the myc family proteins will ultimately be understood in terms of the protein:protein and protein:DNA interactions in which they participate and the biochemical modifications that regulate their activity. Taken together: (1) the cloning of myc target genes presented here and elsewhere (Prendergast and Cole 1989; Eilers et al., 1991; Yang et al., 1991), (2) identification of max, a bHLHzip pairing partner of the myc family proteins (Blackwood and Eisenman 1990; Prendergast et al., 1991), (3) the in vitro definition of an E box DNA sequence to which they can bind (Blackwell et al., 1990; Kerkhoff et al., 1991; Prendergast and Ziff 1991), and (4) development of cell systems in which effects of conditional myc (chapter 1, Cavalieri and Goldfarb 1987; Eilers et al., 1991; Miner and Wold 1991) and/or conditional max (chapter 3) expression can be studied, comprise a powerful set of tools with which to probe myc function.

This thesis includes a substantial contribution to that progress. In chapter 1 we described a conditional c-myc expression system and defined a physiologic setting where that conditional expression produced clear phenotypic effects on both cell cycle progression and mRNA expression. However, the primary impact of that work was that it forced the conclusion that the methodical cloning of myc target genes was required. During that study, possible myc target genes reported in the literature, such as hsp 70 (Kingston et al., 1984; Kaddurah-Daouk et al., 1987), were tested and proved not to be detectably regulated by myc. The experimental system established in that work also suggested an eventually successful approach to the problem. Finally, this cell system provided the experimental basis for a study of the effects of conditional c-myc on MyoD initiated muscle differentiation (Miner and Wold 1991).

Chapter 2 presents the cloning of myc regulated genes. The frequency with which myc targets were identified among a panel of cDNAs subject to either upregulation or downregulation during G_1 of the first cell cycle following serum stimulated emergence from growth arrest (mid-SR class genes) suggests that up to one-third of all such genes may be myc targets. Although the screening effort was extensive, it was not saturating; statistical considerations suggest that we identified only a fraction of the myc target gene class. We were also able to identify a number of other target genes by screening mid-SR class genes cloned elsewhere whose identities suggested their possible involvement in proliferation. In sum, the myc target gene class has been extended to include several extracellular matrix proteins, one anabolic and one catabolic enzyme, a differentiation marker, several important cell proliferation regulators, and an assortment of unidentified genes (see summary table).

While that cloning effort was in progress, two research groups identified max, a gene encoding several bHLHzip proteins that can form DNA binding oligomers either alone or in a complex with myc family proteins. Chapter 3 incorporated conditional max expression into the experimental paradigm begun in chapter 1. Reexamination of the myc target gene set identified individual members that are cooperatively upregulated by myc and max and members that are regulated in opposite directions by myc and by max. In addition we made the suprising observation that max is a regulator of a specific subset of the immediate early serum response gene class.

A summary model of myc and max function.

Although new discoveries concerning the regulatory mechanisms underlying myc's profound biological activities are being made rapidly, enough has been learned through this and other recent work to develop a working model that relates myc and max biochemistry to their cellular

effects. Figure 1 summarizes what has been learned about various myc and max phenotypes, correlates those phenotypes with relative myc and max concentrations, and so provides the framework for the model. The model presented is the simplest one that can accomodate all current data and makes several readily testable predictions. In this model, the primary determinants of the myc overexpressing phenotype are the absolute level of myc:max heteromers and the ratio of myc:max heteromers to max:max homodimers. In addition to these two considerations, it is easy to imagine modulating influences not incorporated in the model. These include competitive interactions with other CACGTG-binding proteins such as USF/MLTF and TFII-i, the possible consequences of biochemical modifications to myc and max, the possibilities for higher-order myc and max containing complexes, and the potential combinatorial interactions between max:max or myc:max containing complexes and other DNA bound regulators. While they must be recognized, these potential influences are not included because the data do not require them and because their consideration at this stage renders the model space unmanageable. Tests of the most simple model will indicate which of these are of greater or lesser importance.

Myc does not function in homotypic complexes. Several lines of evidence suggest that myc proteins cannot form functional homodimers. In vitro, homodimers can only be formed from c-terminal protein fragments, and even then only at very high monomer concentrations (appendix A. See also Blackwell et al., 1990; Kerkhoff et al., 1991; Prendergast and Ziff 1991). Models of the protein:protein dimerization specificity inherent in the leucine repeat motif suggest that the myc leucine repeat may specifically prevent homodimer formation (appendix B, see also O'Shea et al., 1991; O'Shea et al., 1992). In addition, there is evidence from immunoprecipitation analysis that c-myc is quantitatively associated with max in proliferating cells in which max

is apparently present in stoichiometric excess over c-myc (Blackwood et al., 1992).

The role of max. In addition to enabling formation of myc:max heteromers, max proteins may have a function(s) independent of myc proteins. The predicted structure of the max leucine repeat is consistent with favorable homodimer formation (appendix B). In vitro, max homodimers form readily (appendix A; Blackwood and Eisenman 1990; Prendergast and Ziff 1991; Makela et al., 1992), and in our experiments, conditional max expression activated a distinct subset of the immediate early serum response in a setting where there was relatively little c-myc protein present (chapter 3). In addition, max proteins are relatively stable (Blackwood et al., 1992). This stability implies that even low max mRNA levels can support the accumulation of significanct levels of max protein

The importance of the myc:max ratio. Together with the profile of genes regulated by myc and/or max, the data summarized above support a model in which the relative balance of myc and max dictates functionally different outcomes. That difference is apparent at the molecular level, where: (1) distinct but overlapping sets of target genes respond (chapter 3), and (2) where the same gene is a target of both regulatory complexes, regulation can be in opposite directions, e.g., tenascin and c-myc itself (chapters 2 and 3). At the cell biological level, the difference in outcomes is expressed through the distinct phenotypes of max and myc:max overexpressing cells in normal tissue culture growth (chapter 3), ras cotransformation experiments (Prendergast and Ziff 1991; Makela et al., 1992), and cell cycle progression experiments (chapter 3).

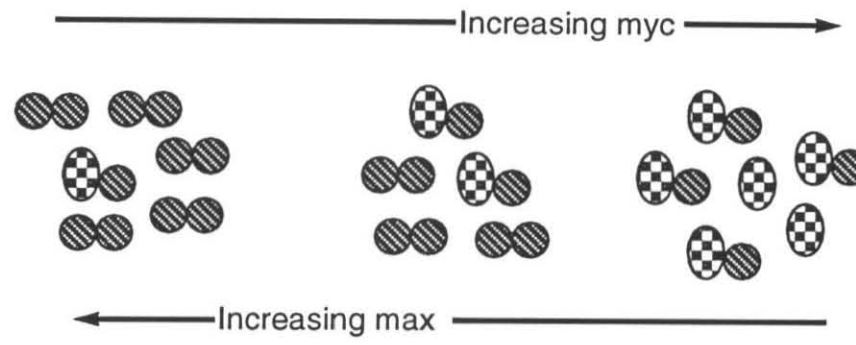
c-Myc, either in a complex with max or in complexes with analogous but as yet unidentified proteins, regulates the expression of a subset of the mid-SR gene class. The work presented here (chapter 3) also

showed that highly elevated max regulates a distinct set of genes that partially overlaps the myc target gene class. In considering the consequences of either myc or max overexpression, it is important that max proteins can form homodimers whereas myc proteins cannot. They therefore have an asymmetric relationship with respect to homotypic and heterotypic complex formation. In particular, myc overexpression leads to overexpression of myc:max heteromers because, in all cases examined thus far, there is always some max present to function as a myc partner. Equally important, we postulate, is the expectation that myc overexpression will lead to downward titration of max:max homodimers. In contrast, max overexpression leads to overexpression of max:max homodimers and alters the ratio of max:max homodimers to myc:max heteromers; however, it cannot titrate nonexistent myc:myc homodimers. As additional myc:max and max:max regulatory targets become known, a primary cause of their pattern of expression will turn out to be understood as the consequence of myc:max overactivity coupled with max:max underactivity.

Activated ras cotransformation assays with c-myc, with forms of max, or with the two genes together suggest that, up to a certain concentration, max and max-9 cooperate with myc in the assay, but at higher concentrations become inhibitory (Prendergast et al., 1991; Makela et al., 1992). The inference is that myc:max heteromers are the positive agent in the ras cotransformations. Max can become limiting for formation of myc:max heteromeric complexes when myc is overexpressed; however, vast excesses of max are expected to lead to max:max homodimers that can compete with, and block the function of, myc:max heteromers. To these observations we add the possibility that excess max competes with myc:max heteromers, not just directly for binding at their target E boxes, but also indirectly by regulating a functionally distinct target gene class (chapte 3). The effect on cotransformation and proliferation suggests that the sum total consequences of max:max overactivity are distinctly anti-proliferative.

Key features of this model are: (1) It posits an active role for max beyond formation of, and competition with myc:max heteromers. (2) It suggests that the ratio of myc:max to max:max is functionally important. (3) It discounts the possibility of myc homodimers, and (4) It suggests that a combination of the absolute myc:max and max:max levels, coupled with their ratio, determines the phenotypic read-out. We can make several experimentally accessible predictions based on the model: (1) In the absence of max, myc overexpression should lead to no observable phenotype. (2) Extreme max overexpression, probably obtainable only in an extremely tight conditional system such as the recently reported VP-16 lac repressor fusion system (Baim et al., 1991), should severely inhibit cell proliferation. (3) Extreme overproduction of either max or myc should be rescued by overproduction of the corresponding bHLHzip partner. (4) Finally, as the identities and functions of the myc:max target gene class members and max:max target gene class members are determined, the sum total change in their expression pattern following alterations in myc or max protein levels will explain the cell proliferation and anti differentiation properties of the myc and max family proteins.

Figure 1



	Excess max	excess myc
myc/max transcritpion factor state:	max homodimers predominate	titrate out max homodimers overexpress myc:max heteromers
growth correlate:	facilitated growth arrest inhibition of myc + ras cotransformation	facilitated G1 progression cellular immortalization inhibition of differentiation
target gene class:	a specific subset of immediate early serum response genes	a specific subset of mid serum response genes
gene expression correlate:	induce c-myc induce I-8-1, tenascin	repress c-myc repress I-8-1, tenascin repress collagen(s) repress I-8-41, HSA
	induce c-jun induce fra-1	
		induce prothymosin α induce I-8-10B, LDH induce I-8-29 induce I-8-30B Induce I-8-36, ODC induce I-8-47 induce p53 tumor antigen induce G1 cyclin D2
	repress G1 cyclin D3	

References

Baim, S. B., Labow, M. A., Levine, A. J., and Shenk, T. (1991). A chimeric mammalian transactivator based on the lac repressor that is regulated by temperature and isopropyl beta-D-thiogalctopyranoside. PNAS **88**(12), 5072-5076

Blackwell, T. K., Kretzner, L., Blackwood, E. M., Eisenman, R. N., and Weintraub, H. (1990). Sequence-specific DNA binding by the c-myc protein. Science **250**, 1149-1151.

Blackwood, E. M., and Eisenman, R. N. (1990). Max: A helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with myc. Science **251**, 1211-1217.

Blackwood, E. M., Luscher, B., and Eisenman, R. N. (1992). Myc and max associate in vivo. Genes and Development **6**(1), 71-80.

Cavalieri, F., and Goldfarb, M. (1987). Growth-factor-deprived BALB/c 3T3 murine fibroblasts can enter the S phase after induction of c-myc gene expression. Molecular and cellular biology **7**, 3554-3560.

Eilers, M., Schirm, S., and Bishop, J. M. (1991). The myc protein activates transcription of the alpha-prothymosin gene. European

Kaddurah-Daouk, R., Greene, J. M., Baldwin, A. S., and Kingston, R. E. (1987). Activation and repression of mammalian gene expression by the c-myc protein. Genes and Development **1**, 347-357.

Kerkhoff, E., Bister, K., and Klempnauer, K. H. (1991). Sequence-specific DNA binding by myc proteins. Proceedings of the National Academy of Sciences USA **88**, 4323-4327.

Kingston, R. E., Baldwin, A. S., and Sharp, P. A. (1984). Regulation of heat shock protein 70 gene expression by c-myc. Nature **312**, 280-282.

Makela, T. P., Koskinen, P. J., Vastrik, I., and Alitalo, K. (1992). Alternative forms of max as enhancers or suppressors of myc-ras cotransformation. Science **256**, 373-376.

Miner, J. H., and Wold, B. W. (1991). c-myc inhibition of myoD and myogenin-initiated myogenic differentiation. Molecular and Cellular Biology **11**(5), 2842-2851.

O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991). X-ray structure of the GCN4 leucine zipper, a two stranded, parallel coiled coil. Science **254**(5031), 539-544.

O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1992). Mechanism of specificity in the fos-jun oncoprotein heterodimer. Cell **68**(4), 699-708.

Prendergast, G. C., and Cole, M. D. (1989). Posttranscriptional regulation of cellular gene expression by the c-myc oncogene. Molecular and Cellular Biology **9**, 124-134.

Prendergast, G. C., Lawe, D., and Ziff, E. B. (1991). Association of Myn, the murine homolog of max, with c-myc stimulates methylation-sensitive DNA-binding and ras cotransformation. Cell **65**(3), 395-407.

Prendergast, G. C., and Ziff, E. B. (1991). Methylation-sensitive sequence-specific DNA binding by the c-myc basic region. Science **251**, 186-189.

Yang, B. S., Geddes, T. J., Pogulis, R. J., de Crombrugghe, B., and Freytag, S. O. (1991). Transcriptional suppression of cellular gene expression by c-myc. Molecular and Cellular Biology **11**(4), 2291-2295.

Appendix A

Towards biochemistry with myc and max

Sean V. Tavtigian, Dawn D. Cornelison, and Barbara J. Wold

Two experimental approaches to myc function are near convergence. Reverse genetic pursuit of myc function through structural homology led both to the cloning of max and demonstration of in vitro DNA binding by max homodimers and myc max heterodimers. At the same time, goal directed cDNA library screening has isolated a suite of max homodimer and myc:max heterodimer target genes. We should soon be able to begin analysis of their function at the molecular biology / biochemistry level.

As a first step in this direction, we have uncovered pH sensitive DNA binding by max homodimers and myc max heterodimers (figure 1), and demonstrated specific binding to a CACGTG E-box in the LDH promoter (figure 2). Using antibodies provided by Elizabeth Blackwood and Bob Eisenman, we have also been able to immunoprecipitate LEmax-9 from nuclear extracts of zinc stimulated {Balb/c 3T3 LEmax-9} 31-4 cells (figure 3).

Several technical points follow from the gelshifts. The translation extracts contained variable amounts of USF/MLTF, identified by its specific binding to the probe and relatively low mobility. We detected two myc:max heteromeric species. The higher mobility species increased its mobility relative to max homodimers with decreasing pH, so that at pH 7.0 the two complexes comigrate. Finally, the USF/MLTF, max homodimer, and myc:max heteromer species were all effeciently competed by unlabelled MM oligo in gelshifts of either the MM or LDH oligos.

Figure 1. pH sensitivity of DNA binding by max homodimers and myc max heterodimers. Synthetic capped RNAs encoding max[1] and LEΔmyc[2] were synthesized and translated essentially according to the manufacturer's recommendations. 2 µl of mock or LEΔmyc translate or 1 µl of max or {LEΔmyc + max} cotranslate were incubated in a 15 µl binding reaction with 50 fmol of the E_{myc} oligo MM[3]. Reactions were incubated at room temp for 30 minutes, fractionated on 5% (39:1 acrylamide:bis) gels, dried, and subjected to autoradiography.

pH 8: Binding reactions were 50 mM Tris Taps pH 8.0. Gel running buffer was 100 mM Tris Taps pH 8.0.

pH 7.5: Binding reactions were 50 mM Tris Mops pH 7.5. Gel running buffer was 100 mM Tris Mops pH 7.5.

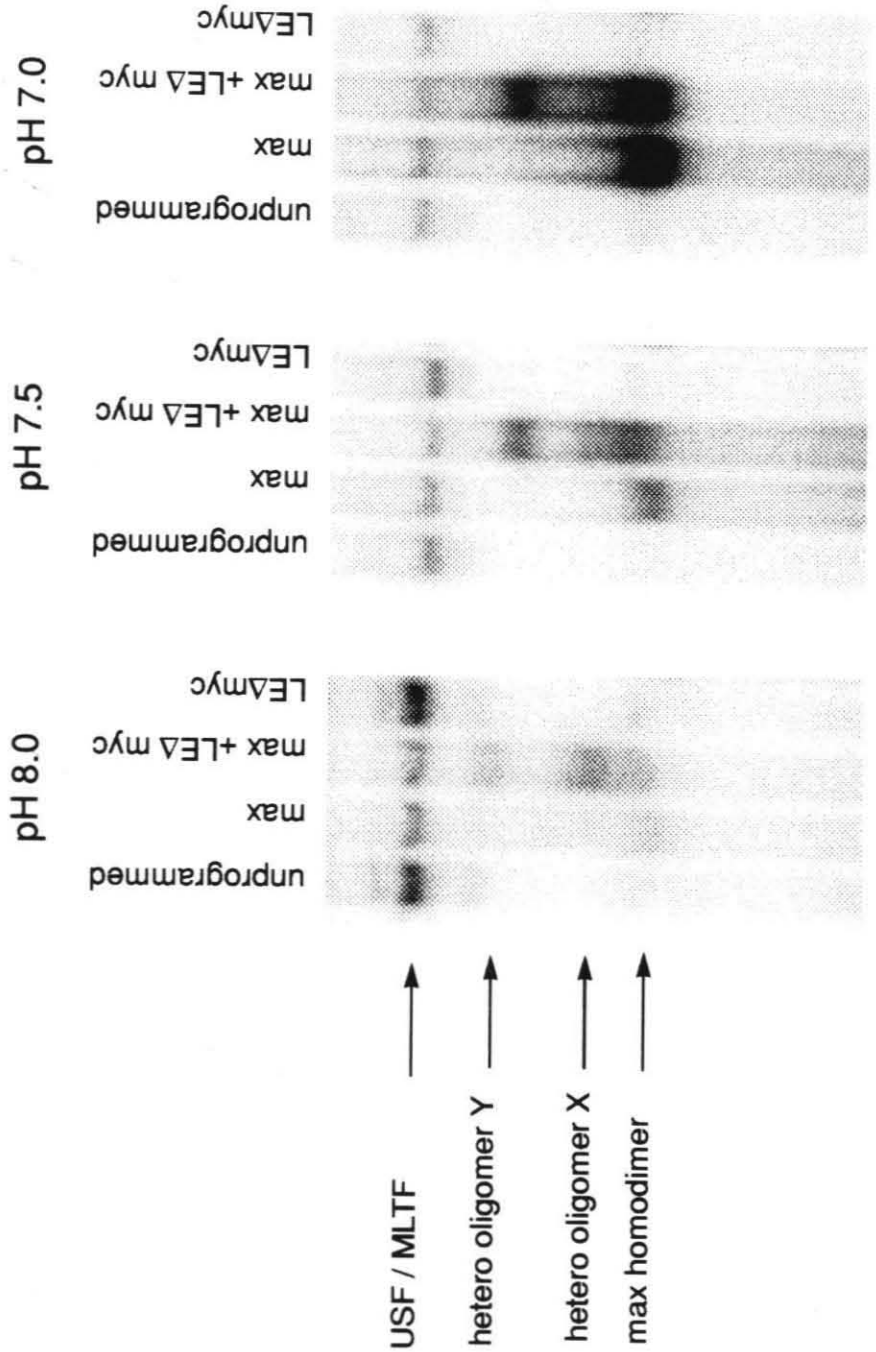
pH 7: Binding reactions were 50 mM Tris Mops pH 7.0. Gel running buffer was 100 mM Tris Mops pH 7.0

In addition to the pH buffer, binding reactions were 50 mM potassium acetate, 5 mM magnesium acetate, 1 mM dithiothreitol, and 5% glycerol. Also, individual 15 µl reactions contained 1 µg of dIdC copolymer and 100 fmol a nonspecific single stranded oligo.

1. The max peptide used in this exercise contained a 1 amino acid substitution, Leu 146 → Phe, well C-terminal to the bHLHzip motif.

2. LEΔmyc is a N-terminal deletion of c-myc that encodes the influenza virus HA1 epitope just upstream of the bHLHzip motif. This deletion is used because full length c-myc is relatively insoluble.

3. MM: TCA GAC <u>CAC GTG</u> GTC GGG TGT TCC


pH 8.0
pH 7.5
pH 7.0
unprogrammed
max
max +LEΔ myc
LEΔmyc
USF / MLTF
hetero oligomer Y
hetero oligomer X
max homodimer

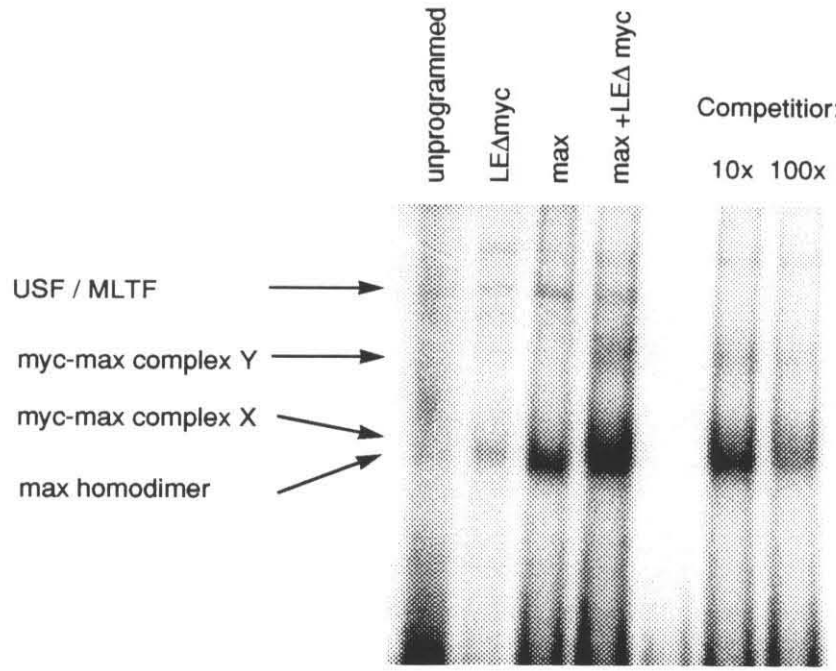


Figure 2. DNA binding by max homodimers and myc max heterodimers at E_{myc} of the lactate dehydrogenase promoter. Synthetic capped RNAs encoding max1 and LEΔmyc2 were synthesized and translated essentially according to the manufacturer's recommendations. 1 µl of mock, LEΔmyc, max or {LEΔmyc + max} cotranslate were incubated in a 15 µl binding reaction with 50 fMol of the LDH promoter derived oligo LDH 4/5[1]. Competitions included 500 fmol (10x) or 5 pmol (100x) of the unlabelled E_{myc} oligo MM (see figure 1). Reactions were incubated at room temp for 30 minutes, fractionated on a 5% (39:1 acrylamide:bis) gel, dried, and subjected to autoradiography. Binding buffer was 50 mM Tris Mops pH 7.0, 50 mM potassium acetate, 5 mM magnesium acetate, 1 mM dithiothreitol, and 5% glycerol. In addition, individual 15 ml reactions contained 1 mg of dIdC copolymer and 100 fmol a nonspecific single stranded oligo. Gel running buffer was 100 mM Tris Mops pH 7.0.

1. LDH 4/5: CCA GCC TAC ACG TGG GTT CCC GCA CGT CCG CTG GGC

Figure 3. Immunoprecipitation of conditionally expressed max from ^{35}S labelled (Balb/c LEmax-9) 34-3 cells. 34-3 cells were seeded on fibronectin coated 15 cm plates and brought to growth arrest by contact inhibition in medium containing 5% calf serum as in materials and methods of chapter 3. At t=0, cells were mock or 20 μM zinc induced. At t=2h, the medium was removed, the cells were washed once with met$^-$ medium, and then refed with 10 ml of met$^-$ medium supplemented with ^{35}S met at 400 μCi/ml, +/- 15 μM zinc. At t=4h, the cells were PBS washed, scraped, and placed on ice. All subsequent work was in the cold room.

The cells were washed with TBS, then swelled on ice for 15 min. in 1 ml of A7.5. The suspension was adjusted to 0.6% NP40, vortexed for 10 sec., and spun at 5,000 RPM for 5 min. to pellet the nuclei. The supernate was set aside as cytoplasmic extract. 150 μl of buffer C was pipetted onto the pellet followed by 15 min. of agitation. This mix was spun at 5,000 RPM for 5 min, and the supernate was set aside as high salt nuclear extract. 150 μl of buffer D was pipetted onto the pellet followed by 15 min. of agitation. This mix was spun at 5,000 RPM for 5 min, and the supernate was set aside as detergent nuclear extract.

The extracts were spun for 1 min. to pellet any remaining debris. Nuclear extracts were adjusted to 1 ml with LS-W2. Extracts were precleared with 10 μl of a 50% slurry of BSA blocked protein A sepharose (Pharmacia), and then incubated with 4 μl of anti max serum (a kind gift of Elizabeth Blackwood and Bob Eisenman), 4 μl of a mixture of 4 anti myc mAbs (including H60C37, the mAb used for the Western blots of chapter 1, provided by Ray Koski at Amgen), and 24 μl of BSA blocked protein A sepharose with agitation for 1 hour. Immunoprecipitates were washed 4 times with LS-W1, boiled in reducing loading dye, fractionated on 17% (170:1 acrylamide:bis) gels, dried, and subjected to fluorography.

Buffers:

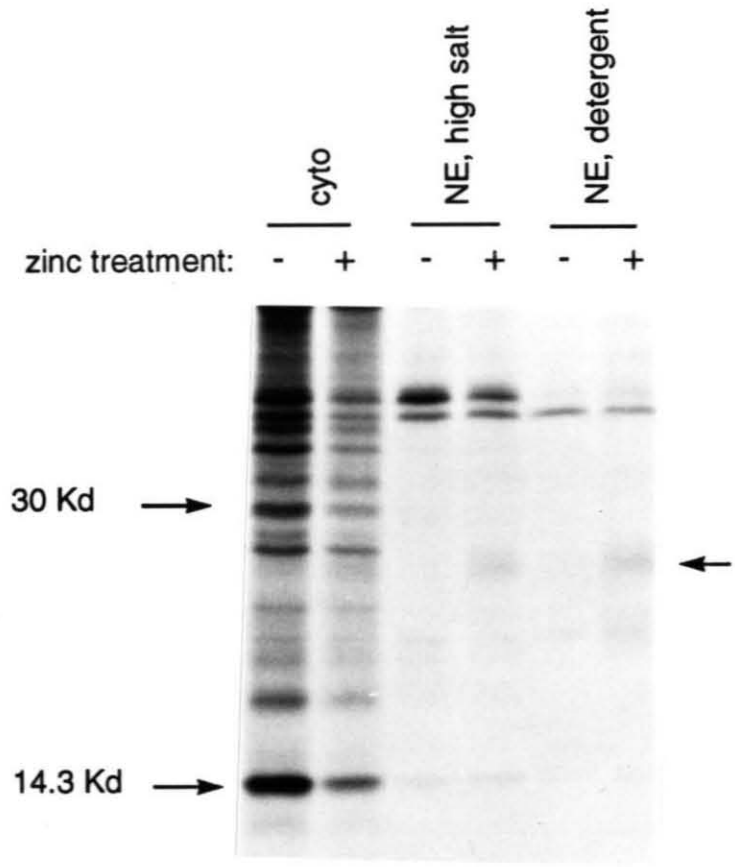
A7.5: 20 mM Tris Mops pH7.5, 10 mM KCl, 1 mM DTT, 0.1 mM EDTA, 0.1mM EGTA, 1.0 mM PMSF, 2.6 µg/ml aprotinin, 1.2 µg/ml antipain, 0.4 µg/ml chymostatin, and 1 µg/ml leupeptin.

C: 25 mM Tris Taps pH 8.0, 400 mM NaCl, 25% glycerol, 1 mM DTT, 1 mM EDTA, 1 mM EGTA, 1 mM PMSF, 2.6 µg/ml aprotinin, 1.2 µg/ml antipain, 0.4 µg/ml chymostatin, and 1 µg/ml leupeptin.

D: 25 mM Tris Taps pH 8.0, 400 mM NaCl, 25% glycerol, 0.5% DOC, 0,5% SDS, 1 mM DTT, 1 mM EDTA, 1 mM EGTA, 1 mM PMSF, 2.6 µg/ml aprotinin, 1.2 µg/ml antipain, 0.4 µg/ml chymostatin, and 1 µg/ml leupeptin.

LS-W1: 1x PBS, 1% NP-40, 10% glycerol, 0.2 mM PMSF, 1.3 µg/ml aprotinin, 0.6 µg/ml antipain, 0.2 µg/ml chymostatin, and 0.5 µg/ml leupeptin.

LS-W2: 0.4x PBS, 1% NP-40, 10% glycerol, 0.2 mM PMSF, 2.6 µg/ml aprotinin, 1.2 µg/ml antipain, 0.4 µg/ml chymostatin, and 1 µg/ml leupeptin.


cyto
NE, high salt
NE, detergent
zinc treatment: - + - + - +
30 Kd
14.3 Kd

Appendix B

Dimerization specificity of the leucine repeat motif

Sean V. Tavtigian

Introduction

Three related groups of DNA binding transcription factors are currently subjects of intense analysis at both functional and structural levels. These are: (1) The basic region - leucine repeat [bzip] proteins, including GCN4, c-fos, c-jun, C/EBP, etc. (2) The basic region - helix-loop-helix [b-hlh] proteins, including myo D1, myogenin, E12, mash1, the asc proteins, etc. (3) The basic region - helix-loop-helix - leucine repeat [bhz] proteins, including c-myc, max/myn, USF/MLTF, TFE-3, AP-4, etc. These groups of proteins share the characteristics that DNA binding is mediated by a basic region which is immediately adjacent to the n-terminus of whichever motif drives protein oligomerization, protein oligomerization is required for DNA binding, and the oligomerization domains are amphipathic α-helices. Also of interest, each group contains members which appear to function as homodimers (or oligomers) [GCN4, E12, USF/MLTF] in vivo as well as members which preferentially (perhaps obligatorily) function as heterodimers (or oligomers)[c-fos, myogenin, c-myc].

In the case of the bzip proteins, the leucine zipper is both sufficient to mediate homo and hetero dimerization and determines pairing specificity. Likewise, the helix-loop-helix motif of the bHLH proteins appears to mediate oligomerization and provide pairing specificity. However, the situation for the bHLHzip proteins may be more complicated. For instance, some work suggests that myc proteins can dimerize via their HLH motif alone[6], whereas AP-4 requires both its HLH and one of its leucine repeat motifs to dimerize, but becomes promiscuous with respect to pairing with bHLH proteins following deletion of its leucine repeats.[5]

While the structure of a dimerized or oligomerized HLH motif is as yet unknown, O'Shea et al. have recently solved a crystal structure of a dimerized GCN4 leucine repeat.[1] As expected, the structure is a parallel two stranded coiled coil. On the basis of their sequence

similarity, the leucine repeats of the other bzip proteins, as well as the leucine repeats of the bHLHzip proteins, should share the same basic structure. Consequently, the GCN4 leucine repeat structure can serve as a starting point to assess/rationalize the observed pairing specificities of other bzip proteins and perhaps help clarify the same issue with respect to the bHLHzip proteins.

The GCN4 Structure

Prior to solution of its crystal structure it was already known that the GCN4 leucine repeat dimerizes as a parallel two stranded coiled coil. Such structures share a series of characteristics[4]:

1. The amino acid sequence is arranged as a heptad repeat [a b c d e f g] wherein the a and d positions are predominantly aliphatic and the g and e positions tend to be charged.

2. The helical repeat length is approximately 3.64 amino acids.

3. The hydrogen bonds parallel to the helix axis tend to be slightly shorter on the hydrophobic (a,d) face of the helix.

4. The hydrophobic face forms the interface between the two strands of the coiled coil. Characteristics 2 and 3 constrain the structure into a shallow left handed supercoil.

5. The crossing angle of the two helices varies between 0 and 30 degrees.

6. The residues at positions a and d of one helix pack "knobs into holes" into the corresponding pocket formed by the a,d,e, and g residues of the other helix.

7. Consequently, Van der Waals interactions along the hydrophobic interface and salt bridges forming between charged residues at the g and e positions provide most of the specificity and stability of the structure.

GCN4 leucine repeat crystals were obtained from a synthetic 33 amino acid peptide corresponding to amino acids 249 to 281 of the native protein. A single heavy metal derivitive was used to obtain initial phase angles. A description of the x-ray diffraction data is given in a single table.[1] Statistical data are presented as overall data to 1.8 A, not by resolution shell. The data set is reasonably complete, with 92% of the native and 87% of the derivitive data used. Reflections were measured twice resulting in respectable R_{sym} values of 0.043 and 0.038. (I'm quoting data for the native structure first and derivitive structure second.) The derivitive seems quite good in that the unit cells dimensions are almost identical and R_{iso} is a respectable 0.20.

On the not so good side, the initial phase angles were ambiguous because only one heavy metal derivitive was used. In addition, the initial electron density map was continuous for 29 amino acids of one chain but discontinuous for the other chain. Consequently, a considerable amount of computer modeling and mathematical manipulation were used to generate the structure. The initial electron density map was solvent flattened and then 29 residues of one chain of the coil were traced. The second chain of the coil was generated by rotating a polyalanine model of the first chain around the noncrystalographic twofold symmetry axis. From this point, the model was used to calculate new phase data and then the electron density map was subjected to eight rounds of twofold averaging, solvent flattening, and phase resolution. At this point the authors had two different models which differed in the register of the sequence along the electron density map. The models were refined against the x-ray data; the one which gave a lower crystallographic R value was then further refined.

After addition of water molecules, the final structure contained 31 of the 33 residues in the peptide. The final crystallographic R value was 0.179 from 6 to 1.8 A resolution; the rms bond angle deviation was 2.5 degrees from ideal and the rms bond length deviation was 0.018 A

from ideal. On the one hand, these statistics suggest that the structure is essentially correct, potentially containing only small errors in sidechain placement. On the other hand, the second chain of the coil was initially generated by a modelling step which assumed the twofold symmetry of the structure, yet the final structure contains a significant assymetry (see later).

Superficial analysis of the GCN4 structure and modeling of c-fos and c-jun homodimers, fos/jun heterodimers, c-myc and max homodimers, and myc/max heterodimers

I found several types of analysis to be informative:

1. Helical net diagrams of the knobs-into-holes packing of each (potential) structure. In each such diagram given, the following methodology is employed:

a. Polarity is n-term at the bottom and c-term at the top

b. structures are distorted to make them fit into a spreadsheet grid. This alters neither prediction of the knobs-into-holes packing around the A and D residues nor prediction of salt bridge interactions at the G and E positions

c. The sequence of one polypeptide chain is given in BLACK. The sense of this chain is such that residue sidechain extend out from the page. (Actually, the [predominantly] leucines at position D extend out nearly perpendicular to the page. The other residues are at various angles to the page.

d. From the GCN4 structure it is further apparent that, in the helical net representation, BLACK CHAIN sidechains at position G point slightly down with respect to the helix axis. BLACK CHAIN sidechains at position E point slightly up with respect to the helix axis.

e. The sequence of the other polypeptide chain is given in RED and residue positions are referred to as x'. In the sense of its orientation, this chain is a mirror image of the black chain; residue sidechains extend into the page. RED CHAIN sidechains at position E' point slightly down with respect to the helix axis, and RED CHAIN sidechains at position G' point slightly up with respect to the helix axis.

f. With respect to this analysis, the main consequence of the G, E, G', and E' sidechain polarity observed in the GCN4 structure is that positions of potential salt bridges can be predicted from the diagram and then checked in the (model) structures. Residues potential participating in salt bridges are circled in green and the distance between the sidechains in the (model) structures are marked. Residues pairs at "salt bridge positions" which should actually repel each other because of their like charges are circled in red.

2. On the computer, the main chain of the polypeptide plus residue sidechains at positions A, D, A'. and D' ,with their Van der Waals surfaces, were created and photographed. Polarity is such that the N-term is always at the bottom and the c-term is always at the top (the same as in the helical net diagrams.)

a. This view allowed me to look for inconsistencies in packing along the hydrophobic interface. Ideally, the analysis would have included looking at VDW surfaces of the E, G, E', and G' residues, but time and overall complication prevented this.

b. In each case the resultant snapshot is named "x.snap1." For instance, the GCN4 example is named GCN4.snap1 and the fos example is named fos.snap1.

3. On the computer, the main chain of the polypeptide plus <u>charged</u> residue sidechains at positions E, G, E', and G' were created and

photographed. Polarity is such that the N-term is always at the bottom and the c-term is always at the top (the same as in the helical net diagrams.)

a. In each case, Asp and Glu residues are colored dark blue. Lys and Arg residues are colored red.

b. Distances between the ends of potentially salt bridged residue pairs were measured. These distance are indicated on the helical net diagrams.

c. In each case the resultant snapshots is named "x.snap2" and, when done, "x.snap3." The third snapshot was only taken when there was something significantly different about the other side of the (model) structure. For instance, the GCN4 examples are named GCN4.snap2 and GCN4.snap3.

Substitution Methodology

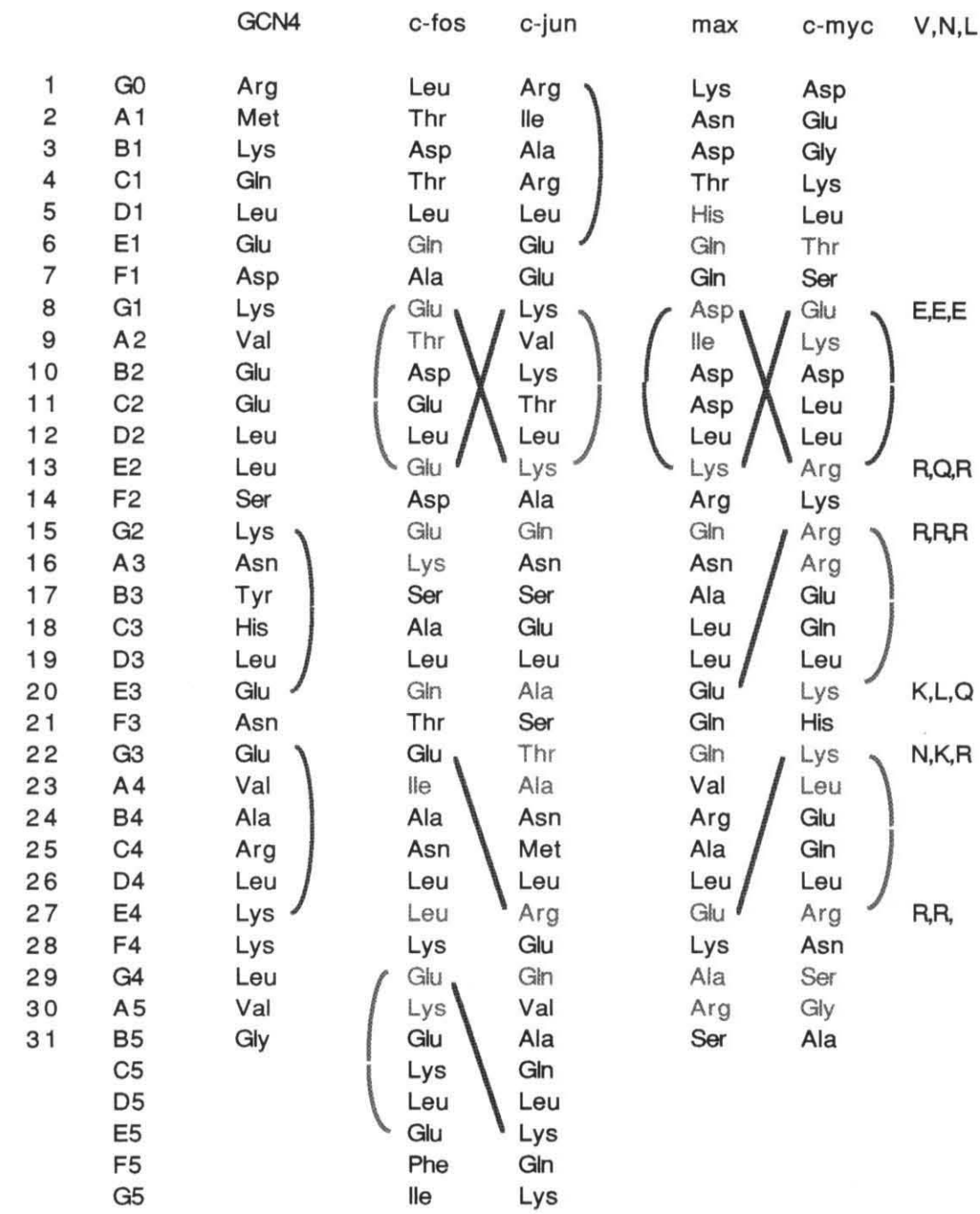
An alignment of the leucine repeat sequences of GCN4, c-fos, c-jun, max, and c-myc is given (figure 1). The polarity is n-term at the top and c-term at the bottom. Helical repeat position assignments (A to G) are given along with residue positions in the dsn2 file.

Substitutions and refinement were as follows:

1. In the sequence alignment, all of the residues marked in red were substituted into the GCN4 dsn2 file using the replace command in SAM.

2. Only residues at positions A, D, E, and G were substituted. Positions B, C, and F were always left as their identity in GCN4. This was done to balance the fact that A, D, E, and G residues participate in dimer formation with the worry that the substituted and refined structures are nonsense (in their fine details) and get steadily worse as the number of substitutions increase.

Figure 1


GCN4
c-fos
c-jun
max
c-myc
V,N,L
1 G0 Arg Leu Arg Lys Asp
2 A1 Met Thr Ile Asn Glu
3 B1 Lys Asp Ala Asp Gly
4 C1 Gln Thr Arg Thr Lys
5 D1 Leu Leu Leu His Leu
6 E1 Glu Gln Glu Gln Thr
7 F1 Asp Ala Glu Gln Ser
8 G1 Lys Glu Lys Asp Glu E,E,E
9 A2 Val Thr Val Ile Lys
10 B2 Glu Asp Lys Asp Asp
11 C2 Glu Glu Thr Asp Leu
12 D2 Leu Leu Leu Leu Leu
13 E2 Leu Glu Lys Lys Arg R,Q,R
14 F2 Ser Asp Ala Arg Lys
15 G2 Lys Glu Gln Gln Arg R,R,R
16 A3 Asn Lys Asn Asn Arg
17 B3 Tyr Ser Ser Ala Glu
18 C3 His Ala Glu Leu Gln
19 D3 Leu Leu Leu Leu Leu
20 E3 Glu Gln Ala Glu Lys K,L,Q
21 F3 Asn Thr Ser Gln His
22 G3 Glu Glu Thr Gln Lys N,K,R
23 A4 Val Ile Ala Val Leu
24 B4 Ala Ala Asn Arg Glu
25 C4 Arg Asn Met Ala Gln
26 D4 Leu Leu Leu Leu Leu
27 E4 Lys Leu Arg Glu Arg R,R,
28 F4 Lys Lys Glu Lys Asn
29 G4 Leu Glu Gln Ala Ser
30 A5 Val Lys Val Arg Gly
31 B5 Gly Glu Ala Ser Ala
C5 Lys Gln
D5 Leu Leu
E5 Glu Lys
F5 Phe Gln
G5 Ile Lys

3. Positions 1 and 2 (G0 and A1) were not substituted because they did not appear to be helical in the GCN4 structure.

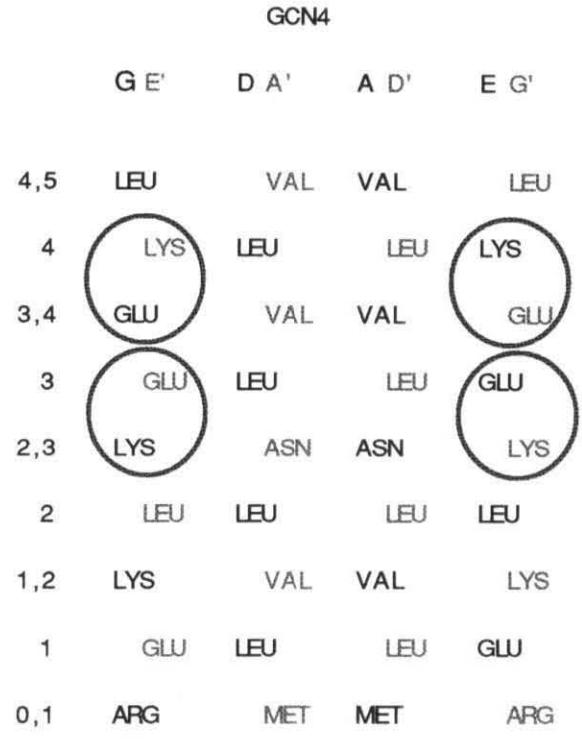
4. After substitution, the A chain was put through 20 cycles of refinement using the commands mode, modl, and refi. Then the B chain was put through 20 cycles of refinement. And then the A chain was put through 20 more cycles of refinement. Finally, the B chain was put through 20 more cycles of refinement. The A and B chains could not be refined simultaneously because the synthetic peptides contained an acetyl group at their n-term.

The GCN4 structure:

Packing along the GCN4 hydrophobic interface conforms almost exactly into the knobs-into-holes model. Leucine residues at the D/D' positions make side to side contacts along their VDW surfaces, and the alternating repeat with Leu at the D/D' positions and Val at the A/A' positions leads to excellent packing with little if any unfilled space (GCN4.snap1). The only irregularity is an A/A' Asn residue (Asn 16) that forms a hydrogen bond within the hydrophobic repeat.

There are four pairs of potential salt bridges between G/E' and G'/E positions (figure 2): Lys 15 to Glu 20 and Glu 22 to Lys 27. The ends of three of these pairs are separated by about 3.5 A in the structure and appear to form salt bridges (GCN4.snap2). The fourth, Lys 15 to Glu 20 on one side of the structure, is forced out of position by the Asn 16 residues (GCN4.snap3). On the other hand, indications from 2D NMR studies are that the solution structure is symmetric.[1]

Figure 2


GCN4
G E'
D A'
A D'
E G'
4,5 LEU VAL VAL LEU
4 LYS LEU LEU LYS
3,4 GLU VAL VAL GLU
3 GLU LEU LEU GLU
2,3 LYS ASN ASN LYS
2 LEU LEU LEU LEU
1,2 LYS VAL VAL LYS
1 GLU LEU LEU GLU
0,1 ARG MET MET ARG

Comparison: c-fos, c-jun, and fos-jun heterodimer model structures

The c-fos structure:

There are no potential salt bridges between the G and E' or E and G' positions. However, there are four pairs of Glu residues which could be involved in like charge repulsion interactions that would destabilize the structure (fos.snap2, figure 3).

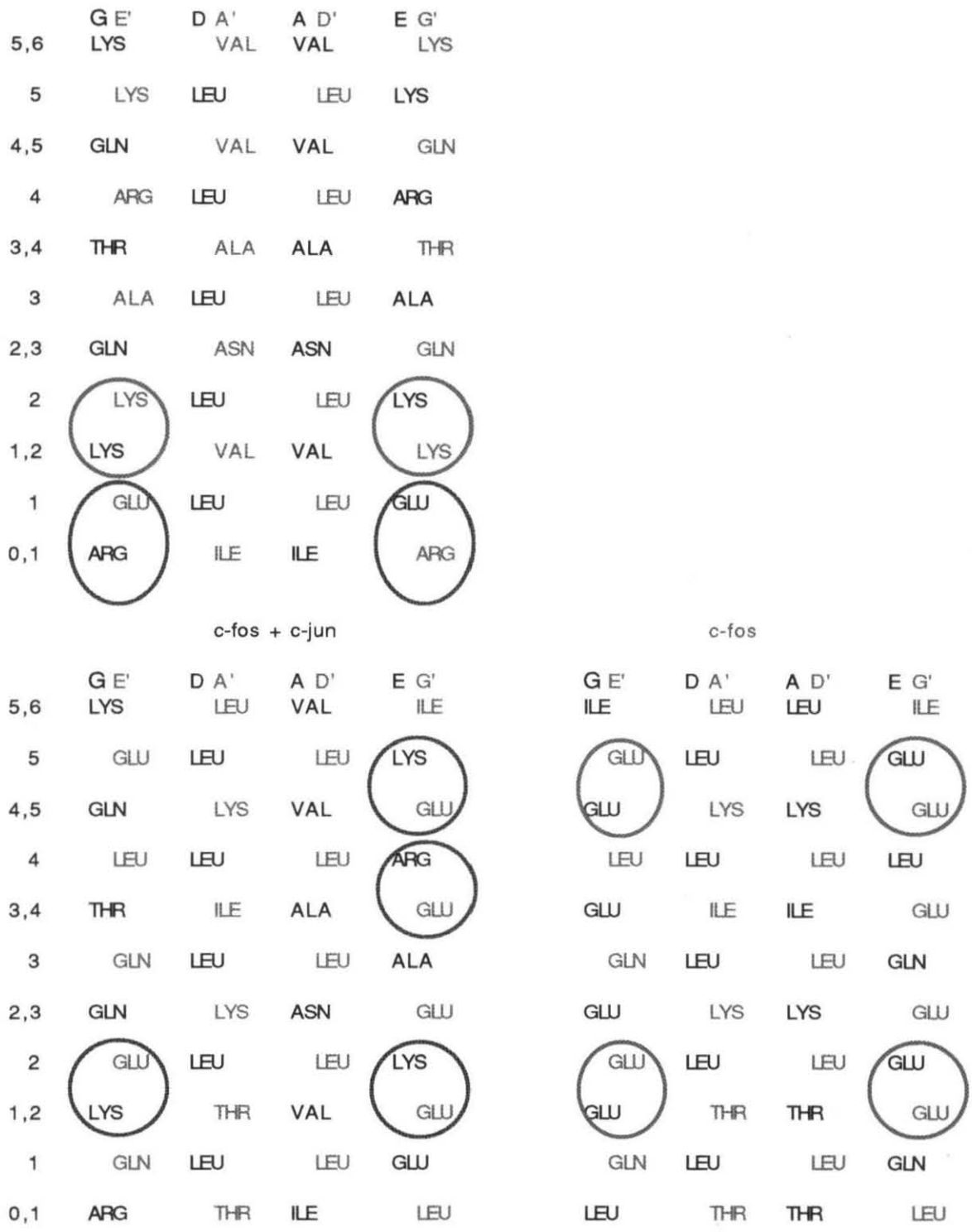
The central hydrophobic interface is quite regular and nicely filled (fos.snap1). However, there are two pairs of Lys residues (positions 16 and 30) buried within the hydrophobic interface. On the one hand, the long aliphatic stretch of the Lys sidechain can pack nicely within the interface. On the other hand, the terminal amino group would seem to be very unfavorable. Interestingly, in the potential structure (fos.snap3) the Lys sidechains extend all the way through the hydrophobic interface to a position where they can form salt bridges with Glu 15 and Glu 29 (fos.snap3) (i.e., from a G'_{n-1} position to an A_n position in the heptad repeat).

The c-jun structure:

From the helical net point of view, there are two potential salt bridges in the structure of the c-jun dimer: Arg 1 to Glu 6 on both sides of the coil. However, this part of the GCN4 structure is not helical, so the physical structure doesn't suggest that particular interaction. However, there are two pairs of Lys residues which could be involved in like charge repulsion interactions that would destabilize the structure (jun.snap2, figure 3).

Like GCN4, c-jun has an Asn residue at an A/A' position (Asn 16) that can form a hydrogen bond within the hydrophobic interface. Beyond that, there is an Ala residue at an A/A' position (Ala 23) which does

Figure 3


c-jun
G E' D A' A D' E G'
5,6 LYS VAL VAL LYS
5 LYS LEU LEU LYS
4,5 GLN VAL VAL GLN
4 ARG LEU LEU ARG
3,4 THR ALA ALA THR
3 ALA LEU LEU ALA
2,3 GLN ASN ASN GLN
2 LYS LEU LEU LYS
1,2 LYS VAL VAL LYS
1 GLU LEU LEU GLU
0,1 ARG ILE ILE ARG
c-fos + c-jun
G E' D A' A D' E G'
5,6 LYS LEU VAL ILE
5 GLU LEU LEU LYS
4,5 GLN LYS VAL GLU
4 LEU LEU LEU ARG
3,4 THR ILE ALA GLU
3 GLN LEU LEU ALA
2,3 GLN LYS ASN GLU
2 GLU LEU LEU LYS
1,2 LYS THR VAL GLU
1 GLN LEU LEU GLU
0,1 ARG THR ILE LEU
c-fos
G E' D A' A D' E G'
ILE LEU LEU ILE
GLU LEU LEU GLU
GLU LYS LYS GLU
LEU LEU LEU LEU
GLU ILE ILE GLU
GLN LEU LEU GLN
GLU LYS LYS GLU
GLU LEU LEU GLU
GLU THR THR GLU
GLN LEU LEU GLN
LEU THR THR LEU

not efficiently fill the space around it in the hydrophobic repeat (jun.snap1).

The fos-jun heterodimer structure

From the helical net analysis, there are three potential salt bridges in the heterodimer structure (figure 3): (fos Glu 22 to jun Arg 27, fos Glu 8 to jun Lys 13, and fos Glu 13 to jun Lys 8). In the model structure, the first of these is separated by only 3.3A, while the other two are 7.5 to 8 A apart. There are no potential charge repulsion interactions that would destabilize the structure

The central hydrophobic repeat is better than that of c-jun in the sense that jun Ala 23 is now next to fos Ile 23, filling the space between them much better. A/A' position 16 now has a Lys and an Asn residue. The Lys residue cannot form a salt bridge with a nearby Glu residue, as in c-fos, so this interaction would seem particularly unfavorable.

Comparison: c-myc, max, and myc-max heterodimer model structures

The c-myc structure:

From the helical net point of view, there is one pair of potential salt bridges in the c-myc structure (figure 4): Glu 8 to Arg 13. However, both pairs of residues are separated by 7.9 A in the model structure (myc.snap2). On the other hand, there are two pairs of like charge repulsion interactions that would destabilize the structure: Arg 15 to Lys 20 and Lys 22 to Arg 27.

The central hydrophobic repeat is very irregular. There are three charged residues buried at A/A' positions: Glu2, Lys 9, and Arg 16. Furthermore, the Lys residues do not seem to fill their region of the hydrophobic interface very well (myc.snap1).

The max structure:

From the helical net point of view, there is again one pair of potential salt bridges in this structure: Asp 8 to Lys 13. However, both pairs of residues are separated by 8.9 A in the model structure (max.snap2, figure 4). There are no potential like charge repulsive interactions.

The central hydrophobic repeat is considerably more regular than that of c-myc. Although there are two pairs of buried Asn residues (Asn 2 and Asn 16) which may form hydrogen bonds as in GCN4, there are no buried charged residues. The hydrophobic repeat also appears well packed (max.snap1).

The myc-max heterodimer structure:

The helical net analysis predicts four possible salt bridges in the myc-max heterodimer structure (figure 4): myc Glu 8 to max Lys 13, max Asp 8 to myc Arg 13, myc Arg 15 to max Glu 20, and myc Lys 25 to max Glu 27. In the model strucure these charge pairs are separated by 7.9, 9.0, 5.2, and 6.2 A, respectively (mycmax.snap2 and mycmax.snap3). There are no potential like charge repulsive interactions.

The central hydrophobic repeat contains three buried charges residues contributed by c-myc: Glu 2, Lys 9, and Arg 16. Max also contributes two Asn residues: Asn 2 and Asn 16. On the other hand, the VDW surface packing seems better than that of c-myc (mycmax.snap1)

Figure 4 C-MYC:

	G E'	D A'	A D'	E G'
4,5	SER	GLY	GLY	SER
4	ARG	LEU	LEU	ARG
3,4	LYS	LEU	LEU	LYS
3	LYS	LEU	LEU	LYS
2,3	ARG	ARG	ARG	ARG
2	ARG	LEU	LEU	ARG
1,2	GLU	LYS	LYS	GLU
1	THR	LEU	LEU	THR
0,1	ASP	GLU	GLU	ASP
0	GLN	ILE	ILE	GLN
0,-1	TYR	ILE	ILE	TYR

MYC + MAX

	G E'	D A'	A D'	E G'
4,5	SER	ARG	GLY	ALA
4	GLU	LEU	LEU	ARG
3,4	LYS	VAL	LEU	GLN
3	GLU	LEU	LEU	LYS
2,3	ARG	ASN	ARG	GLN
2	LYS	LEU	LEU	ARG
1,2	GLU	ILE	LYS	ASP
1	GLN	LEU	HIS	THR
0,1	ASP	ASN	GLU	LYS
0	ARG	ILE	MET	GLN
0,-1	TYR	ILE	ILE	TYR

MAX

G E'	D A'	A D'	E G'
ALA	ARG	ARG	ALA
GLU	LEU	LEU	GLU
GLN	VAL	VAL	GLN
GLU	LEU	LEU	GLU
GLN	ASN	ASN	GLN
LYS	LEU	LEU	LYS
ASP	ILE	ILE	ASP
GLN	HIS	HIS	GLN
LYS	ASN	ASN	LYS
ARG	MET	MET	ARG
TYR	ILE	ILE	TYR

Discussion

The main questions are whether or not the structural analysis presented can help rationalize pairing specificities observed in vitro and in vivo, and whether or not the modeling has any predictive value.

The answer to the first question is clearly yes. In the Fos/Jun world the order of observed binding affinity is fos+jun > jun+jun > fos+fos. Helical net analysis suggests that fos/jun has three potential charge bridges, one of which looks good in the model. And the central hydrophobic repeat looks better than either that of jun or fos homodimers. C-jun homodimers have no potential charge bridges and two potential like charge repulsive interactions. But this is more favorable than c-fos which had 4 potential like charge repulsive interactions at the E/G' and G/E' positions as well as four Lys residues buried in its hydrophobic repeat.

The situation with c-myc and max is qualitatively similar. The observed pairing affinities are myc+max > max+max > myc+myc. Helical net analysis suggests that myc+max has four potential salt bridges, none of which look particularly good in the model. VDW surfaces fill most of the space within the central hydrophobic repeat, but there are several buried charges. This seems better than max+max which has two potential salt bridges, neither of which looks particularly good, plus two pairs of Asn residues within the central hydrophobic repeat which presumably form hydrogen bonds as in GCN4. Finally, myc+myc has two potential salt bridges between E/G' or E'/G positions, but it also has four potential like charge repulsive interactions at similar positions. There appears to be one gap in the packing of the central hydrophobic repeat; furthermore, that repeat is not particularly hydrophobic - it includes 6 charged residues!

Predictive issues are entirely experimental. For instance, could one design and synthesize peptides which would pair more efficiently with a given bzip protein than its natural pairing partner? Or could

one use synthetic peptides to clone as yet unknown bzip or bHLHzip proteins? Time will tell.

References

1. O'Shea, E.K., Klemm, J.D., Kim, P.S., and Alber, T. 1991. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science 254: 539-544.

2. Blackwood, E.M. and Eisenman, R.N. 1990. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with myc. Science 251: 1211-1217.

3. Prendergast, G.C., Lawe, D., and Ziff, E.B. 1991. Association of Myn, the murine homolog of max, with c-myc stimulates methylation-sensitive DNA-binding and ras cotransformation. Cell 65(3): 395-407

4. Cohen, C. and Parry, D.A.D. 1990. α-helical coiled coils and bundles: how to design an α-helical protein. Proteins 7: 1-15.

5. Hu, Y-F., Luscher, B., Admon, A., Mermod, N., and Tjian, R. 1990. Transcription factor AP-4 contains multiple dimerization domains that regulate dimer specificity. Genes and Development 4(10): 1741-1752.

6. Kerkhoff, E., Bister, K., and Klempenauer, K-H. 1991. Sequence-specific DNA binding by myc proteins. PNAS 88: 4323-4327.

7. Mouse c-fos and c-jun sequences were from Genbank.

Appendix C

Restriction map comparison programs

Sean V. Tavtigian

Computerized restriction map comparison

An important practical goal in medium to large scale cloning exercises is often to reduce labor by eliminating redundant clones as early in the process as possible. In our work, this goal was accomplished in three steps. The first step was accomplished at the level of plaque screens, and the third step required sequence analysis (chapter 2). The second step involved computerized comparison of restriction maps generated from the 3' ends of the selected clones.

The programs used in this step were written in APL. The first program, REGRESSION, generates standard curves from gel profiles of molecular weight markers. The second program, MAP, uses those standard curves to interpolate the molecular weights of restriction fragments generated by partial restriction digests, and formats the restriction map data for the map comparison programs.

The map comparison programs are MAPCOMPARE and MULTICOMPARE. Given an ordered array of restriction fragment lengths generated by digests of any number of clones by a single enzyme, MAPCOMPARE compares the maps and determines which are likely to represent clones of the same gene. In essence, the program calculates the variance between the logically comparable fragments for each pair of clone. MULTICOMPARE runs MAPCOMPARE as a subroutine. It allows comparison of the data from any number of restriction digests of the same set of clones, increasing the probability of correctly identifying redundant clones.

```
          rE←REGRESSION ;rA;rB;rC;rD
[1]   'Enter vector of molecular weights'
[2]   rA←10⍎⍞
[3]   'Enter vector of gel positions'
[4]   rB←⍞
[5]   →(Out,Reg)[1+(⍴rA)=⍴rB]
[6]   Out: 'The # molecular weights ≠ # gel positions'
[7]   →0
[8]   Reg: 'What order regression would you like to do?'
[9]   rC←⍞
[10]  rD←((⍴rA),(rC+1))⍴rD←rA⌹rB∘.*(0,⍳rC)
[11]  ⍞←'Coefficients are ',⍕rD[1;]
[12]  rE←rD[1;]
[13]  ' '
[14]  ⍞←'Calcd mws',6 0⍕10*+/rD×rB∘.*(0,⍳rC)
[15]  ⍞←'real mws ',6 0⍕10*rA


          MAP ;mA;mB;mC;mD;N
[1]   'What are the coefficients of the standard curve?'
[2]   mA←⍞
[3]   'How many digests are you going to analyze?'
[4]   mB←⍞
[5]   cMap←(mB,10)⍴0
[6]   cName←(mB,6)⍴'cDNA  '
[7]   N←1
[8]   'What is the name of cDNA # ',⍕N
[9]   cName[N;]←6↑mC←⍞
[10]  'Enter the gel band positions of ',cName[N;],' in ascending order'
[11]  cMap[N;]←10↑⍎mD←⍞
[12]  N←N+1
[13]  →8 14[1+N>mB]
[14]  cMap←(cMap>0)×10*+/(((⍴cMap),⍴mA)⍴mA)×cMap∘.*(¯1+⍳⍴mA)
[15]  ⍞←cName,6 0⍕cMap
[16]  cFrag←(cMap-¯1⌽cMap)
[17]  ⍞←cName,6 0⍕cFrag←(mB,9)↑(cFrag>0)×cFrag
```

```
     ∇ Compmat←MAPCOMPARE ;cB;cD;cE;N
[1]   'Enter vector of cDNA names'
[2]   cNames←¯1⍎((⍴cNames)+0 2)↑cNames←⍞
[3]   'Enter matrix of ordered cDNA fragment sizes'
[4]   cB←⍞
[5]   →(Branch1,Branch2) [1+1↑(⍴ cNames)= ⍴ cB]
[6]   Branch1: '# of cDNA names ≠ # of cDNA fragment lists'
[7]   →Finished
[8]   Branch2: 'How many leading fragment sizes are of no comparative value?'
[9]   cB←(0,⍞)↓cB
[10]  'What is the minimum # of valid fragment comparisons required for this'
[11]  'analysis ?'
[12]  N←⍳1+⍞
[13]  delNames←,(~×/cB[;N]≠0)/[1]cNames          ⍝Store names to-be-deleted
[14]  cNames←(×/cB[;N]≠0)/[1]cNames              ⍝Delete names
[15]  cB←(×/cB[;N]≠0)/[1]cB                      ⍝Delete fragment data
[16]  Compmat←((1↑⍴cB),1↑⍴cB)⍴0                  ⍝Format comparison matrix
[17]  N←1
[18]  Loop: cD←(¯1++/cB[N;]>0)↑cB[N;]            ⍝Take all but last frag length
[19]  cD←((1↑⍴cB),⍴cD)⍴cD                        ⍝and make data into a matrix
[20]  cD←cD-(⍴cD)↑cB                             ⍝and subtract fragment lengths
[21]  cE←.5×((⍴cD)⍴+\cD[N;])+(⍴ cD)↑ +\cB        ⍝Make ave. res. site pos. matrix
[22]  Compmat[N;]←+/(cD*2)÷cE                    ⍝Insert sum-of "variance"
[23]  Compmat[N;]←Compmat[N;]÷¯1↑⍴cD             ⍝and divide by # of comparisons
[24]  ⍞←'This is fun ',⍕N
[25]  N←N+1
[26]  →(Loop,Out) [1+N>1↑⍴ cNames]
[27]  Out:cNames[((+/cB>0)≤2)/⍳1↑⍴cB;1]←'⊛'      ⍝Flag names corr. to short maps
[28]  ⍞←'Remaining storage in workspace: ',⍕ (⎕WA)-16000
[29]  cB←cD←cE←0                                 ⍝Get rid of unneeded variables
[30]  'Do you want to print the matrix of comparison data (y/n) ?'
[31]  →(Branch4,Branch3)[1+(1↑⍞)∈'yY']
[32]  Branch3: PRINTCOMP
[33]  Branch4: 'Set the maximum value for the comparison statistic (1-10)'
[34]  Compmat←0<(⍉Compmat)+Compmat←Compmat≤⌈⍞     ⍝Make reflected logical matrix
[35]  Finished: 'The map comparison is finished'
     ∇
```

```
     ∇MULTICOMPARE ;mcA;mcB;mcC;mcD;J
[1]   'WARNING:  the order of cDNAs within the name and fragment lists must'
[2]   'be identical [between comparisons] or else the program will generate'
[3]   'nonsense comparison data'
[4]   ' '
[5]   'How many digests are being "compared" ?'
[6]   mcA←⍞
[7]   J←1
[8]   Loop1: 'Calling MAPCOMPARE for analysis of digest #',⍕J
[9]   →(Branch1,Branch2)[1+J>1]
[10]  Branch1: mcB←MAPCOMPARE                    ⍝Enter comparison matrix
[11]  mcD←cNames[;1]≠' '                         ⍝Note flagged cDNA names
[12]  →Continue
[13]  Branch2: mcB←mcB×(⍴mcB)↑MAPCOMPARE         ⍝"And-compare" comp. matrices
[14]  mcD←mcD×cNames[;1]≠' '                     ⍝"And-compare" flagged names
[15]  Continue: →(Loop1,Output)[1+mcA<J←J+1]
[16]  Output: 'The output is ready.  Print output or screen output (p/s) ?'
[17]  →(Screen,Print)[1+(1↑⍞)∈'Pp']
[18]  Print: ⎕PRSELECT                           ⍝Direct ouput to printer
[19]  Screen: mcC←(×/mcB)                        ⍝Find "all-match" cDNAs
[20]  delNames←delNames,,mcC/[1]cNames           ⍝and record their names
[21]  cNames[;1]←' ⊕'[1+mcD]                     ⍝Re-flag cDNA names
[22]  cNames←(~mcC)/[1]cNames                    ⍝Delete "all-match" cDNA names
[23]  mcB←(~mcC)/[1]mcB←(~mcC)/[2]mcB            ⍝Delete corresponding data
[24]  J←1
[25]  'cDNA:          Potential matches'         ⍝Print list of matching cDNAs
[26]  Loop2:  ⎕←cNames[J;],':  ',,cNames[mcB[J;]/⍳(1↑⍴mcB);]
[27]  →(Loop2,Out)[1+(1↑⍴mcB)<J←J+1]
[28]  Out: ' '                                   ⍝Print list of noncompared cDNAs
[29]  'cDNAs deleted for lack of comparable fragments in any digest:'
[30]  ' '
[31]  ⎕←delNames
[32]  ⎕←'Comparison output completed'
     ∇
```

Appendix D

Partial cDNA sequnces of mid-SR clones

Sean V. Tavtigian and Julianne M. Rogers

```
LOCUS       I8-01       243 BP DS-DNA              UPDATED   11/30/90
DEFINITION  - Tenascin
ACCESSION   - D90343
LENGTH      - 1320
SOURCE      - Balb/c 3T3
BASE COUNT       61 A     60 C     69 G     53 T      0
ORIGIN      ?
    I8-01.5 Length: 243   Dec  4, 1990 - 04:47 PM   Check: 1657 ..
  1 GCCCGGGTGT AGAGTACCGC GTCAGTGTGA TTGCCATGAA GGGATTCGAA GAAAGCGATC
 61 CAGTCTCGGG GACTCTAATC ACAGCTCTGG ATGGTCCATC TGGTCTTCTG ATAGCCAACA
121 TCACAGACTC AGAAGCCTTG GCCATGTGGC AGCCAGCCAT TGCCACTGTG GACAGTTATG
181 TCATCTCCTA CACAGGGGAG AGAGTGCCAG AAGTTACACG CACAGTGTCT GGAAATACAG
241 TGG
//
```

```
LOCUS       I8-06       259 BP DS-DNA              UPDATED   08/25/90
DEFINITION  - melanoma growth stimulatory activity
ACCESSION   - X12510
LENGTH      - 1175
SOURCE      - Balb/c 3T3
BASE COUNT       88 A     51 C     68 G     52 T      0 OTHER
ORIGIN      ?
    I8-06.5  Length: 305  September 19, 1990  15:38  Check: 9991  ..
  1  ggtaccggat cgaattcgaa cccttcggTA ACTCCAAAAA TTAATCCCAA
 51  ATTGATCGCT AATTTGGAGG TGATCCCTGC AGGTCCACAG TGCCCTACGG
101  TGGAAGTCAT AGCTAAACTG AAAAACCAGA AGGAGGTCTG TCTGGATCCA
151  GAAGCTCCTG CGATAAAGAA AATCATTCAG AAAATATTGG GCAGTGACAA
201  AAAGAAAGCT AAGCGGAATG CACTCGCAGT GGAAAGAACG GCCAGTGTTC
251  AATAGAAAGA TTTCTGAGGA CTCTGACCCC AGTGAAGAGA AGAAGGGCTG
301  ATTCT
```

LOCUS I8-09 288 BP DS-DNA UPDATED 12/04/90

DEFINITION - not identified

ACCESSION -

LENGTH - 2030

SOURCE - Balb/c 3T3

BASE COUNT 63 A 90 C 65 G 70 T 0 OTHER

ORIGIN ?

I8-09.5 Length: 288 Dec 4, 1990 - 04:48 PM Check: 3160 ..

1 CCTTCGCCCA AACCCAAACC ATATCTGGGG CGGTGGGATT CTCTACTGCC AAGCACCCAT

61 CCTTGGAAGA AAGTTGTTGG CTGATATGCT GATGCTTCCT TGACGTCACC AGAGAGTCCT

121 CCTCTAGCCA CCAAATATGG CCCCGTCCAT CCTCAATTAC ATACACTCGG GGCCTCCCCA

181 GCTGCCATAC CACTGGCGCC ACTCTTGAGG GTGGCTGCTG GGTCACACAC TGAGGTCTTC

241 CACATCCATA TCATCAAGTT CTGATGGTGG TTCAGGTCTT AGCAAGAG

//

LOCUS I8-10A 298 BP DS-DNA UPDATED 08/25/90

DEFINITION - mouse T1 protein

ACCESSION - M24843

LENGTH - 2300

SOURCE - Balb/c 3T3

BASE COUNT 91 A 67 C 59 G 66 T 0 OTHER

ORIGIN ?

I8-10A 5' Length: 298 Sep 4, 1990 - 10:00 AM Check: 9682 ..

1 ACATAAAAAG CCCGCCAAGC TGCAATATCC CTGATTATTT GATGTACTCG ACAGTACGTG

61 GATCAGATAA AAATTTCAAG ATAACGTGTC CAACAATTGA CCTGTATAAT TGGACAGCAC

121 CTGTTCAGTG GTTTAAGAAC TGCAAAGCTC TCCAAGAGCC AAGGTTCAGG GCACACAGGT

181 CCTACTTGTT CATTGACAAC GTGACTCATG ATGATGAAGG TGACTACACT TGTCAATTCA

241 CACACGCGGA GAATGGAACC AACTACATCG TGACGGCCAC CAG

//

LOCUS I8-10B 273 BP DS-DNA UPDATED 08/16/90

DEFINITION -lactate dehydrogenase

ACCESSION - M27554

LENGTH - 780

SOURCE - Balb/c 3T3

BASE COUNT 62 A 67 C 77 G 67 T 0 OTHER

ORIGIN ?

I8-10b.5 Length: 343 September 19, 1990 15:40 Check: 2593 ..

```
  1  ggtaccggat cgaattcgaa cccttcggAA TCCTTAGGCG GGTGCATCCC
 51  TAAACCACCA TGATTAAGGG TCTCTATGGA ATCAATGAGG ATGTCTTCCT
101  CAGTGTCCCA TGTATCCTGG GACAAAATGG AATCTCGGAT GTTGTGAAGG
151  TGACACTGAC TCCTGAGGAA GAGGCCCGCC TGAAGAAGAG CGCAGACACC
201  CTCTGGGGAA TCCAGAAGGA GCTGCAGTTC TAAAGTCTTC CCCGTGTCCT
251  AGCACTTCAC TGTCCAGGCT GCAGCAGGGC TTCTAGGCAG ACCACACCCT
301  TCTCGTCTGA GCTGTGGTTA GTACAGTGGT GTTGAGATGG TGT
```

//

LOCUS I8-16 295 BP DS-DNA UPDATED 11/02/90

DEFINITION - cytoplasmic (A-X) actin

ACCESSION - J04181

LENGTH - 1420

SOURCE - Balb/c 3T3

BASE COUNT 62 A 88 C 80 G 65 T 0 OTHER

ORIGIN ?

I8-16.5 Length: 295 Nov 1, 1990 - 04:39 PM Check: 5972 ..

```
  1 TTCGCATCTA CGAGGGCTAT GCTCTCCCTC ACGCCATCCT GCGTCTGGAC CTGGCTGGCC
 61 GGGACCTGAC AGACTACCTC ATGAAGATCC TGACCGAGCG TGGCTACAGC TTCACCACCA
121 CAGCTGAGAG GGAAATCGTG CGTGACATCA AAGAGAAGCT GTGCTATGTT GCTCTAGACT
181 TCGAGCAGGA GATGGCCACT GCCGCATCCT CTTCCTCCCT GGAGAAGAGC TATGAGCTGC
241 CTGACGGCCA GGTCATCACT ATTGGAACGA GCGGTTCCGA TGCCTGAGGC TCTTT
```

//

```
LOCUS       I8-23       227 BP DS-DNA              UPDATED   02/07/91
DEFINITION  - not identified
ACCESSION   -
LENGTH      - 660
SOURCE      - Balb/c 3T3
BASE COUNT       58 A     34 C     58 G     77 T      0 OTHER
ORIGIN      ?
    I8-23.5 Length: 227   Feb  6, 1991 - 11:15 AM   Check: 6022 ..
  1 CCCCTTCGAT ACAATTTAAT TCCAGATTTC TTTTTGGGGG GCTTAAGTGG CATCAAATTT
 61 TGGTATATTT CTGTCTTATG TTAAAGAAAT ATATTGCTAA AATGTCAGGG AGCGCTAATG
121 TCAGCTGTCA AGGACTAGAT TTACTTTCGC AGGATATGGA GCTCAGTGAG CTGAGGTGGT
181 GAGGTGTATG TGATCTGTGG GAATTCTGCC ATTTAACACA GGAGGTT
//
```

```
LOCUS       I8-29       321 BP DS-DNA              UPDATED   11/30/90
DEFINITION  - not identified*
ACCESSION   -
LENGTH      - 2850
SOURCE      - Balb/c 3T3
BASE COUNT       78 A     83 C     91 G     69 T      0 OTHER
ORIGIN      ?
    I8-29.5 Length: 321   Dec  4, 1990 - 04:50 PM   Check: 4063 ..
  1 TTCGAAAGGA TCTATGACAA TGAGCTGCCC CTGGGCAACA CCTCAGCTTC CAGACAAGCT
 61 TCGAGAGGCA GACATTGTGG TCTTAGGCTC ACCTAAGCCA GAAGAGATTC CTGCTGCGTG
121 GATTCCATCA GGAACCACCA TTCTCAATTG CTTCCATGAC TTCTTGTCAG GGAAGCTGAG
181 CGGAGGTTCT CCTGGGGTCC CCGTGGACAA ACTCATTGCA GAAGAGAGCG TGAGTCTCCT
241 TGCTGCGGCT CTGCGCATTC AGAACATGGT GAGCAGCGGC AGGAGGTGGC TCAGAGAGCA
301 GCAACATCGA AGATGGCGAC T
//
```

LOCUS I8-30B 229 BP DS-DNA UPDATED 12/02/90

DEFINITION - not identified

ACCESSION -

LENGTH - 800

SOURCE - Balb/c 3T3

BASE COUNT 64 A 58 C 63 G 44 T 0 OTHER

ORIGIN ?

I8-30B.5 Length: 229 Dec 4, 1990 - 04:51 PM Check: 7396 ..

1 TGCTTCCACC TTGGATCCCT GGGCTTTCTG ACCCCCTTCA ACTTTGAGAA CTTCCAATCC

61 CAAGTGAATC AGGTGATAGA GGGGAATGCC GCTGTCATCC TGAGGAGCCG GCTGAAGGTC

121 AGGGTGGTGA AAGAGCCCAG AGACAAGAAG ACAGCCATCC ACAACGGCCT CAGTGAAAAT

181 GGTCTGGACA CAGAGGGCGG AAAACAGGCT ATGCAATACC AGGTCTTAA

//

LOCUS I8-32 282 BP DS-DNA UPDATED 12/02/90

DEFINITION - liver thiol transferase

ACCESSION - M31453

LENGTH - 1280

SOURCE - Balb/c 3T3

BASE COUNT 76 A 78 C 62 G 66 T 0 OTHER

ORIGIN ?

I8-32.5 Length: 282 Dec 4, 1990 - 04:52 PM Check: 5237 ..

1 ACCATTACTT TCAACTGCAC GTTCCTCCCT GGAGAAGCTG CAGCCTGTCA GCATGGCTCA

61 GGAGTTTGTG AACTGCAAGA TCCAGTCTGG GAAGGTGGTC GTGTTCATCA AGCCCACCTG

121 CCCCTACTGC AGAAAGACCC AAGAAATCCT CAGTCAACTG CCTTTCAAAC AAGGTCTTCT

181 GGAGTTTGTG GACATCACAG CCACTAACAA CACCAGTGCG ATTCAAGATT ATTTACAACA

241 GCTCACCGGA GCGAGAACAG TTCCTCGGGT CTTCATAGGT AA

//

LOCUS I8-33 296 BP DS-DNA UPDATED 10/25/90

DEFINITION - glyceraldehyde-3-phosphate dehydrogenase

ACCESSION - M32599

LENGTH - 600

SOURCE - Balb/c 3T3

BASE COUNT 62 A 89 C 83 G 62 T 0 OTHER

ORIGIN ?

I8-33.5 Length: 296 Oct 24, 1990 - 12:16 PM Check: 3693 ..

1 TTCGGGGCCC AGAACATCAT CCCTGCATCC ACTGGTGCTG CCAAGGCTGT GGGCAAGGTC

61 ATCCCAGAGC TGAACGGGAA GCTCACTGGC ATGGCCTTCC GTGTTCCTAC CCCCAATGTG

121 TCCGTCGTGG ATCTGACGTG CCGCCTGGAG AAACCTGCCA AGTATGATGA CATCAAGAAG

181 GTGGTGAAGC AGGCATCTGA GGGCCCACTG AAGGGCATCT TGGGCTACAC TGAGGACCAG

241 GTTGTCTCCT GCGACTTCAA CAGCAACTCC CACTCTTCCA CCTTCGATGC GGGGCT

//

LOCUS I8-34 282 BP DS-DNA UPDATED 10/25/90

DEFINITION - not identified

ACCESSION -

LENGTH - 2900

SOURCE - Balb/c 3T3

BASE COUNT 74 A 63 C 62 G 83 T 0 OTHER

ORIGIN ?

I8-34.5 Length: 282 Oct 24, 1990 - 12:16 PM Check: 4092 ..

1 CTTCGTTTGT AATCCCTTCA CAGTCCCAGG TTTAGTGAAA AACTGCTGTA AACACGGGGG

61 ACACAGTTTA ACAATGCAAC TTTTAATGAC TGTTTTCTTT TTCCTTAACT TACTAATAGT

121 TTGTGGATCT GATAAGCAAG GGTGTGTGGT TGAAGAAAAC CTCTGTGGTG GGCTTAATCA

181 GTCACTACTA CTATGCAAAC CCTAAACCGG CACCCTGGTG ACCGGGGGCA TTCGTATAAG

241 AAAAGCATTG TGTGTGACTC TGTGTCCACT CAGACCGACC CC

//

LOCUS I8-36 343 BP DS-DNA UPDATED 10/25/90

DEFINITION - ornithine decarboxylase

ACCESSION - M10624

LENGTH - 1070

SOURCE - Balb/c 3T3

BASE COUNT 100 A 82 C 79 G 82 T 0 OTHER

ORIGIN ?

I8-36.5 Length: 343 Oct 24, 1990 - 12:17 PM Check: 8105 ..

1 CTTCGGCTTC TCCCAGTGTA ATCAACCCAG CTCTGGACAA GTACTTCCCA TCAGACTCTG

61 GAGTGAGAAT CATAGCTGAG CCAGGCAGAT ACTATGTCGC ATCAGCTTTC ACGCTTGCAG

121 TCAACATCAT TGCCAAAAAA ACCGTGTGGA AGGAGCAGCC CGGCTCTGAC GATGAAGATG

181 AGTCAAATGA ACAAACCTTC ATGTATTATG TGAATGATGG AGTATATGGA TCATTTAACT

241 GCATTCTTTA TGATCATGCC CATGTGAAGG CCCTGCTGCA GAAGAGACCC AAGCCAGACG

301 AGAAGTATTA CTCATCCAGC ATCTGGGGAC CAACATGTGA TGG

//

LOCUS I8-41 302 BP DS-DNA UPDATED 11/02/90

DEFINITION - heat stable antigen

ACCESSION - M58661

LENGTH - 1850

SOURCE - Balb/c 3T3

BASE COUNT 61 A 95 C 76 G 70 T 0 OTHER

ORIGIN ?

I8-41.5 Length: 302 Nov 1, 1990 - 04:40 PM Check: 803 ..

1 TTCGCGAACA TCTAGAGAGT CGCGCCGCGC GCCGACGGAG CGGACATGGG CAGAGCGATG

61 GTGGCCAGGC TAGGGCTGGG GTTGCTGCTT CTGGCACTGC TCCTACCCAC GCAGATTTAC

121 TGCAACCAAA CATCTGTTGC ACCGTTTCCC GGTAACCAGA ATATTTCTGC TTCCCCAAAT

181 CCAAGTAACG CTACCACCAG AGGGGGTGGC AGCTCCCTGC AGTCCACAGC TGGTCTCCTG

241 GCTCTCTCTC TCTCTCTTCT ACATCTCTAC TGTTAGAGAC TCAGGCAGGA AACGTCTCTA

301 CT

//

LOCUS I8-47 251 BP DS-DNA UPDATED 10/19/90

DEFINITION - not identified

ACCESSION -

LENGTH - 620

SOURCE - Balb/c 3T3

BASE COUNT 56 A 50 C 55 G 90 T 0 OTHER

ORIGIN ?

I8-47 5' Length: 251 Oct 20, 1990 - 09:15 PM Check: 5395 ..

1 ACATTGCTTT TGAAATTTCA AAAGTCATTT TAATACACAC TCACTGAGCT GCCCATGTGT

61 TTTTAATTGT GCTACCATAG AAGTTCTGTG GAAGATTTTA TTATTGTTGG TTTGATTTTG

121 GTTTTTTAGA GGTGGGGTCT TGCTATGTAG CCCAGGCTGG TCTTGAACTT GCTATGTAGC

181 CTGACTGGTC TTGAACCCCT GACCTCCTGT GTCCCAAGTG CTAGGATCAC AAAAGACATC

241 TCCATGCCTG G

//

LOCUS R8-01 304 BP DS-DNA UPDATED 11/02/90

DEFINITION - Alpha-1 (3) collagen

ACCESSION - X06700

LENGTH - 1460

SOURCE - Balb/c 3T3

BASE COUNT 74 A 76 C 97 G 57 T 0 OTHER

ORIGIN ?

R8-1.5 Length: 304 Nov 1, 1990 - 04:41 PM Check: 4582 ..

1 TTCGTGGAGC TCGAGGTGCT CCGGGTCCCC AAGGTCCACG AGGTGACAAA GGTGAAACTG

61 GTGAACGTGG CTCTAATGGC ATCAAAGGAC ATCGAGGATT CCCTGGCAAT CCAGGTCCTC

121 CAGGTTCTCC TGGTGCTGCT GGTCACCAGG GTGCAATTGG TAGTCCAGGA CCTGCAGGTC

181 CCAGAGGACC AGTTGGACCA CATGGACCTC CTGGAAAAGA TGGAACAAGT GGGCATCCAG

241 GTCCTATTGG ACCACCAGGT CCTAGAGGAA ACAGAGGTGA AAGAGGATCT GAGGGCTCGC

301 CAGG

//

LOCUS R8-09 283 BP DS-DNA UPDATED 02/07/91

DEFINITION - Elongation factor 1

ACCESSION -

LENGTH - 1800

SOURCE - Balb/c 3T3

BASE COUNT 89 A 63 C 77 G 54 T 0 OTHER

ORIGIN ?

R8-9.5 Length: 283 Feb 13, 1991 - 11:21 PM Check: 8560 ..

```
  1 GGGAGCGGGT TTGCCGTCAG AACGCAGGTG TTGTGAAAAC CACCGCTAAT TCAAAGCAAA
 61 AATGGGAAAG GAAAAGACTC ACATCAACAT CGTCGTAATC GGACACGTAG ATTCCGGCAA
121 GTCCACCACA ACCGGCCACC TGATCTACAA ATGTGGTGGA ATCGACAAGC AACCATCGAA
181 AAGTTTGAGA AGGAGGCTGC TGAGATGGGA AAGGGCTCCT TCAAGTACGC CTGGGTCTTA
241 GACAAACTGA AGCTGAGCGT GAGGCTGGTA TCACTATTCG AAT
```

//

LOCUS R8-26 315 BP DS-DNA UPDATED 11/02/90

DEFINITION - Alpha-1 (6) collagen

ACCESSION - J04598

LENGTH - 2300

SOURCE - Balb/c 3T3

BASE COUNT 83 A 75 C 94 G 63 T 0 OTHER

ORIGIN ?

R8-26.5 Length: 315 Nov 1, 1990 - 04:41 PM Check: 1755 ..

```
  1 TTCGGGAGAC CCAGGAGAGG ATAACAACGA CATTTCACCC CGTGGGGTCA AAGGGGCAAA
 61 GGGATACCGA GGCCCAGAAG GACCCCAGGG ACCTCCAGGA CATGTGGGAC CACCTGGGCC
121 AGATGAGTGT GAGATCCTGG ATATCATCAT GAAAATGTGC TCCTGCTGTG AGTGCACATG
181 TGGACCCATT GACATCCTCT TCGTGCTGGA CAGCTCGGAG AGCATTGGCC TACAGAACTT
241 TGAGATTGCC AAGGACTTCA TCATCAAGGT CATTGACCGG TTGAGCAAGG ATGAGCTGGT
301 CAAATTTGAG CCAGG
```

//