

MINI
REVIEW

Discordance between gene regulation *in vitro* and *in vivo*

Richard N. Kitsis¹ and Leslie A. Leinwand^{1,2}

¹Departments of Medicine and ²Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, New York

Mammalian gene expression can be regulated by cell type, developmental stage, and numerous external stimuli. Regulation is brought about by a complex interaction of cis-acting DNA sequences and trans-acting factors. Considerable effort has been expended to identify and characterize regulatory DNA sequences and to understand the basic mechanisms by which they exert their actions. Traditionally, the transcriptional activities of putative promoters and regulatory sequences have been measured by transfections into cultured cells of various DNA constructs consisting of these sequences fused to a reporter gene and subsequent quantitation of reporter gene expression. This approach has provided a great deal of information concerning the sequences mediating the expression of genes in various cell types in response to physiologic perturbations such as hormones—and, to a lesser extent, during specific developmental stages, such as the fusion of myoblasts to form myotubes.

Given the complexity of biological processes *in vivo*, however, a question arises whether regulatory elements identified in cultured cells accurately explain the regulation of gene expression *in vivo*. This becomes particularly important in considering the use of immortalized cell lines or primary cultures, whose patterns of gene expression often differ significantly from those of the intact tissue.

One way to evaluate the relevance of *in vitro* transfection studies to gene regulation *in vivo* is to create transgenic animals using constructs that have also been tested in cultured cells. This

approach requires the generation of several independent mouse lines for each construct under study, an expensive and time-consuming undertaking compared with transfections into cultured cells. One alternative to transgenic mice that may approximate *in vivo* gene regulation more closely than cell culture transfections is the injection of DNA sequences linked to reporters into cardiac and skeletal muscle tissue *in vivo* (Kitsis et al., 1991). However, this gene transfer technique appears to be limited to striated muscle cells (Wolff et al., 1991).

Regardless of whether gene regulation is studied *in vitro* or *in vivo*, the state of the foreign DNA may influence the regulation of its expression. With transient transfections into cultured cells or injections of DNA into muscle *in vivo* (Wolff, et al., 1990), the construct being tested exists in an episomal state. The geometric configuration of the regulatory sequences in the construct may differ markedly from its natural configuration in the endogenous locus. This, in turn, may influence regulation of the expression of the construct. In addition, an episome will not be subject to regulatory information that may reside in the chromatin configuration. On the other hand, with stably transfected cells and transgenic animals, the construct undergoes integration into random chromosomal sites whose chromatin configuration may differ from that of the endogenous locus. In fact, the importance of long-range chromosome effects is supported by the existence of locus control regions (LCR), such as those seen in the β globin gene complex (see Grosfeld, et

al., 1987; Forrester et al., 1987), which act at a distance to confer copy number-dependent, position-independent expression onto a gene. Although it would be optimal to study the regulation of a construct at its endogenous locus, or at least within its LCR, this is usually not possible. Therefore, studies should be performed on multiple independent cell or transgenic lines to separate effects which result from the cis-acting elements under study from those which are due to the position of integration.

In addition to transcriptional regulation, posttranscriptional control plays an important role in the control of mammalian gene expression. Expression of a test construct consisting of promoter and regulatory elements linked to a reporter gene will not reflect posttranscriptional processes that may be relevant to the regulation of the gene under study, since sequences specifying the endogenous pattern of RNA splicing and mRNA stability have been replaced with heterologous sequences. Some investigators have used constructs containing a minigene, rather than a reporter gene, to address some of these issues (for example, see Dente et al., 1988). Though obviously important, the regulatory role played by posttranscriptional events is beyond the scope of the present review.

Although *in vivo* experiments have often yielded results identical to those obtained in cultured cells (for example, see Leask et al., 1990), there are multiple examples of discordance between gene regulation *in vitro* and *in vivo*. The purpose of this review is to discuss some of the examples in which transcriptional regulation conferred by a given cis-acting DNA sequence element differs significantly between cultured cells and *in vivo*. In addition, possible mechanisms for the discordance will be discussed.

Gene regulation *in vitro* and *in vivo*

In a relatively small number of cases, some of the same DNA constructs tested in transfections have also been evaluated in transgenic mice (Table 1). This section will compare results that have been obtained using both approaches. Discordance between gene regulation *in vitro* and *in vivo* can take the form of "promiscuous" expression of the DNA construct in transfections into different cell types with appropriate expression in transgenics. In the case of the liver-

specific α_1 -acid glycoprotein gene (AGP-A), ~1.2 kb of upstream sequence and ~2 kb of 3'-flanking sequence were able to direct transcription with similar efficiencies, both in HeLa cells, where the gene "should not" be expressed, and in Hep3B hepatoma cells, which express the endogenous AGP-A gene. In contrast, in transgenic mice, the same sequences resulted in appropriate liver-restricted expression (Dente et al., 1988). This type of discordance is also exhibited by a human *N-myc* minigene, which was expressed promiscuously following transient and stable transfections into mouse 3T3 cells, where the endogenous gene is not expressed, but was expressed concordantly with the endogenous gene in transgenic animals (Zimmerman et al., 1990). The endogenous rat α -cardiac myosin heavy chain gene is expressed in cardiac but not skeletal muscle. Despite this, ~0.6 kb of 5'-flanking sequence was sufficient to direct transcription in both myoblasts and myotubes of the L₆E₉ rat skeletal muscle cell line (Tsika et al., 1990). This sequence was not transcriptionally active in HeLa cells or 3T3 fibroblasts, demonstrating limited—but not entirely specific—expression of this construct *in vitro*. In contrast, the same upstream sequence directed cardiac-specific expression in transgenic mice (Katz et al., 1992) and following injection into the hearts of adult rats *in vivo* (Kitsis et al., 1991).

Another form of discordance is observed when cis-acting elements are sufficient to direct expression in transfected cells but are inadequate to drive transgene expression in the analogous cell type *in vivo* or in embryonic stem cells. For example, transfection of an α -fetoprotein minigene into hepatoma cells demonstrated the functional equivalence of three enhancers located in the 5'-flanking region (Godbout et al., 1986). In contrast, constructs containing various combinations of these elements differed in their capacities to direct transcription in the livers of transgenic animals (Hammer et al., 1987). Another example is the mouse α_1 -antitrypsin gene. Approximately 0.5–10 kb of 5'-flanking sequence results in expression of a human growth-releasing factor minigene in transiently and stably transfected hepatoma cells and primary hepatocyte cultures, but not in 3T3 or neuroblastoma cells (Montgomery et al., 1990). In contrast, identical constructs lacked sufficient information to be expressed in multiple lines of transgenic animals

Table 1. Discordance between gene regulation in vitro and in vivo

Expression pattern	Gene	Sequences	Construct	In vitro		In vivo		References
				Transfection	Cell type	Tissue		
In vitro—promiscuous and In vivo—appropriate	Human α_1 -acid glycoprotein	~1.2 kb 5' ~2.0 kb 3'	M	T	Hep3B (hepatoma) HeLa	Liver		Dente et al., 1988
	Human N-myc	~3 kb 5' ~3 kb 3'	M	T,S	Transferred gene expressed in multiple cell lines that do not express endogenous N-myc	Brain Kidney		Zimmerman et al., 1990
	Rat α -cardiac myosin heavy chain	~0.6 kb 5'	R (in vitro) M (transgenics) R (injections)	T	1° fetal cardiocytes L6E9 myoblasts L6E9 myotubes	Heart		Tsika et al., 1990 Katz et al., 1992 Kitsis et al., 1991
In vitro—appropriate and In vivo—none, less, or requiring additional sequences	Mouse α -fetoprotein	~6.6 kb 5'	M	T	HepG2 (hepatoma) (constructs with 3 enhancers functionally equivalent)	Liver (constructs with 3 enhancers not functionally equivalent)		Godbout et al., 1986 Hammer et al., 1987
	Mouse α_1 -antitrypsin	~0.5–10 kb 5'	M	T,S	HepG2 (hepatoma) BWTG-3 (hepatoma) SK-N-SH/SY5Y (neuroblastoma) NIH 3T3 (fibroblast)	None		Montgomery et al., 1990 Harris and Krauter, personal communication
	Human HPRT	See reference for details	M	T	CHO Human fibroblasts Mouse L (constructs with or without intronic sequences expressed)	ES (only constructs with intronic sequences expressed)		Reid et al., 1990
	Mouse muscle creatine kinase	~1.3–3.3 kb 5'	R	T	MM14 myotubes (equal expression with 3.3 and 1.3 kb constructs)	Skeletal muscle (3.3 kb 10- to 100-fold more active than 1.3 kb)		Jaynes et al., 1988 Johnson et al., 1989

Abbreviations: ES, embryonic stem cells; HPRT, hypoxanthine phosphoribosyl transferase; M, minigene; R, reporter construct; S, stable transfection; T, transient transfection.

(T. Harris and K. Krauter, personal communication). Discordance between gene expression in somatic and embryonic stem (ES) cells has also been observed with a hypoxanthine phosphoribosyl transferase (HPRT) minigene, which was expressed following transfection into multiple somatic cell types but not into ES cells, which express the endogenous HPRT gene (Reid et al., 1990). Addition of HPRT intronic sequences to the minigene, however, resulted in its expression in ES cells. Gene expression in cultured cells can be quantitatively discordant from that in transgenic animals, even if expression is limited to the appropriate tissues in both systems. For example, ~1.3 kb of mouse muscle

creatine kinase (MCK) upstream sequence was as efficient as a ~3.3 kb construct in driving reporter gene expression in differentiated MM14 cells, a mouse myoblast line (Jaynes et al., 1988). In several lines of transgenic animals, however, the longer construct was 10- to 100-fold more transcriptionally active than the shorter one (Johnson et al., 1989).

Cell culture models for studying gene regulation

Given the discordance between gene regulation in vitro and in vivo, it is worthwhile to examine mechanisms that might explain the phenom-

enon. One place to begin is to compare quantitative and qualitative patterns of endogenous gene expression in cultured cells and intact tissues. For example, many hepatoma cell lines have been shown to express liver-specific genes at very low levels or not at all (Clayton et al., 1985). The H₄A₇C₂ rat hepatoma cell line, for instance, expresses (among others) one liver-specific gene, glutathione S-transferase, but only at 5% of the level seen in the intact liver. Although culture conditions can be manipulated to induce expression of this gene several fold, it remains substantially lower than that observed *in vivo* (Gatmaitan et al., 1983). In addition, culture conditions designed to promote liver-specific gene expression have failed to activate the liver-specific α_1 -antitrypsin gene, which is silent in H₄A₇C₂ cells. Muscle cell lines also exhibit differences in gene expression relative to normal muscle tissue. This is exemplified by the L6J1 myoblast line, which—although capable of fusing into myotubes and initiating a program of muscle-specific gene expression—does not express the entire program of myogenic regulators, since it does not express MyoD (Hinterberger, et al., 1991). Additionally, the normal temporal and quantitative program of myosin heavy chain gene expression does not occur in muscle cell lines (see Weydert et al., 1987). In summary, the phenotype of many cell lines diverges significantly from that of cells in the intact organ.

There are several potential mechanisms for altered gene expression in cell lines. First, the genotype of immortalized cells may be altered, which may affect the expression of specific genes. Second, cells in culture lack many of the cues that trigger physiological and developmental processes. For example, paracrine and mechanical signals from other cell types with which they are in contact in a tissue undoubtedly influence gene expression. Other stimuli that may be different (or lacking) in cell culture relate to the physical substrate on which cells are grown and the media in which they are bathed. Of note, using a chemical and/or a second cell type to model the extracellular matrix and hormonally defined media to replicate the extracellular milieu has been shown to modulate gene expression in cultured cells to mimic more closely that *in vivo* (see Jefferson et al., 1984; Enat et al., 1984). In addition, cultured

cells do not experience the myriad of developmental stimuli to which cells that are formed and grow in an embryo are subjected. It is possible that the absence of this developmental history results in alterations in the subsequent program of gene expression.

Primary cell cultures, which may more accurately reflect gene expression *in vivo*, are an alternative to immortalized cell lines for gene regulation studies. In some cases, such as cardiac myocytes, primary cultures must be used, since there are as yet no permanent cell lines. Unfortunately, gene expression in primary cultures can also be discordant with that in the same cell type in the environment of the intact organ. This can be the result of the cell dissociation procedure, cell culture conditions (see Jefferson et al., 1984), or the absence of the appropriate developmental stimuli, as discussed above. When hepatocytes are disaggregated from the liver and cultured on a plastic substrate in serum-containing media, the cells assume a flat morphology, and liver-specific gene expression decreases dramatically, while the expression of housekeeping genes such as tubulin remains constant (see Clayton and Darnell, 1983; Clayton et al., 1985). Albumin transcription decreases to 50% of baseline after 4 hours and to 10% after 24 hours in culture (Xanthopoulos et al., 1989). Similarly, levels of skeletal α -actin mRNA specifically decline in cultured neonatal cardiac myocytes compared with the intact neonatal heart (Bishopric et al., 1987).

In other cases, the discordance between gene expression *in vitro* and *in vivo* may reflect the phenotype of the cells at the time they are placed in culture. For example, it has been possible to use only fetal and neonatal cardiocytes in transfection studies. The phenotype of these cells differs significantly from that of adult cardiocytes, thereby limiting the applicability of the results of gene regulation experiments in these primary cultures. In support of this idea, it has recently been shown that DNA constructs that are inactive in cultured fetal cardiocytes are active when injected into the adult heart *in vivo* (P. M. Buttrick et al., unpublished data). Despite the limitations of extrapolating results obtained with myocytes from an earlier developmental stage to adults, neonatal cardiac myocyte cultures have been useful for analyzing some complex physiological processes, such

as cellular hypertrophy, in which patterns of gene expression in cultured neonatal cells mimic those in vivo (Waspe et al., 1990).

Bases of discordance

Clues to the molecular basis for discordance between gene expression in cultured cells and the same cell type in the intact tissue have been provided by analyses of trans-acting factors. Steady-state levels of mRNAs encoding several transcription factors have been observed to decrease in cultured liver cells. For example, the sequence-specific DNA binding protein C/EBP is expressed at levels that are at least 10-fold lower in cultured hepatoma cells compared with the intact liver (Friedman et al., 1989). Further, it has been demonstrated that transcription of this gene decreases almost 10-fold within 24 hours of placing hepatocytes into culture (Xanthopoulos et al., 1989). In fact, recent studies suggest that decreases in C/EBP occur in response to activation of cell proliferation, thereby suggesting that expression of high levels of C/EBP might be incompatible with most rapidly growing cell lines (Mischoulon et al., 1992). Another example is provided by repression of the expression of HNF-4, a member of the steroid hormone receptor superfamily, in dedifferentiated hepatomas and in hepatocyte-somatic cell hybrids in which the liver phenotype was extinguished (Kuo et al., 1992). This alteration, in turn, resulted in absence of HNF-1 α , a homeoprotein that requires HNF-4 to transactivate its expression. In these systems, expression of other transcription factors, such as C/EBP, was unchanged, demonstrating that the observed changes in gene expression were not the result of general decreases in the abundance of all transcription factors. The implication of such findings for the study of gene regulation is significant. If certain trans-acting factors are required to achieve appropriate tissue- or developmental stage-specific or stimulus-responsive expression, then gene regulation in cultured cells, which express different levels of these factors compared with the intact tissue, may be discordant with that in vivo. It seems likely that some of the cases of discordance discussed in this review can be attributed to altered expression of trans-acting factors in cultured cells.

In summary, the regulation of many mam-

malian genes is different in cultured cells than in intact tissues. Future studies directed at defining the differences in cellular milieu in these two states will undoubtedly extend our understanding of this discordance and of the underlying basis for gene regulation in vivo.

Acknowledgments

We thank Drs. Kenneth Krauter and Peter Buttrick for helpful discussions.

This work was supported by National Institutes of Health grants HL 02699 (R. N. Kitsis), GM 29090, and HL 37412 (L. A. Leinwand); American Heart Association, New York City Affiliate Grants-in-Aid (R. N. Kitsis and L. A. Leinwand); and the Irma T. Hirschl Scholar Award (L. A. Leinwand).

The costs of publishing this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC Section 1734 solely to indicate this fact.

References

- G. Acsadi, S. Jiao, A. Jani, D. Duke, P. Williams, W. Chong, and J. A. Wolff (1991), *New Biol* 3, 71-81.
- N. H. Bishopric, P. C. Simpson, and C. P. Ordahl (1987), *J Clin Invest* 80, 1194-1199.
- D. F. Clayton, M. Weiss, and J. E. Darnell, Jr. (1985), *Mol Cell Biol* 5, 2633-2641.
- D. F. Clayton and J. E. Darnell (1983), *Mol Cell Biol* 3, 1552-1561.
- L. Dente, U. Ruther, M. Tripodi, E. Wagner, and R. Cortese (1988), *Genes Dev* 2, 259-266.
- R. Enat, D. M. Jefferson, N. Ruiz-Opazo, Z. Gatmaitan, L. A. Leinwand, and L. M. Reid (1984), *Proc Natl Acad Sci USA* 81, 1411-1415.
- W. C. Forrester, S. Takegawa, T. Papayannopoulou, G. Stamatoyannopoulos, and M. Groudine (1987), *Nucleic Acids Res* 15, 10159-10177.
- A. D. Friedman, W. H. Landschultz, and S. L. McKnight (1989), *Genes Dev* 3, 1314-1322.
- Z. Gatmaitan, D. Jefferson, N. Ruiz-Opazo, L. Biempica, I. Arias, G. Dudas, L. Leinwand, and L. M. Reid (1983), *J Cell Biol* 97, 1179-1190.
- R. Godbout, R. Ingram, and S. M. Tilghman (1986), *Mol Cell Biol* 6, 477-487.
- F. Grosveld, G. Blom van Assendelft, G. Greaves, and G. Kollias (1987), *Cell* 51, 975-985.
- R. H. Hammer, R. Krumlauf, S. A. Cramper, R. L. Brinster, and S. M. Tilghman (1987), *Science* 235, 53-58.
- T. J. Hinterberger, D. A. Sassoon, S. J. Rhodes, and S. F. Konieczny (1991), *Dev Biol* 147, 144-156.
- J. B. Jaynes, J. E. Johnson, J. N. Bushin, C. L. Gartside, and S. D. Hauschka (1988), *Mol Cell Biol* 8, 62-70.

- D. M. Jefferson, D. F. Clayton, J. E. Darnell, Jr., and L. M. Reid (1984), *Mol Cell Biol* 4, 1929–1934.
- J. E. Johnson, B. J. Wold, and S. D. Hauschka (1989), *Mol Cell Biol* 9, 3393–3399.
- E. B. Katz, M. E. Steinhelper, J. B. Delcarpio, A. J. Daud, W. C. Claycomb, and L. J. Field (1992), *Am J Physiol* 31, H1867–H1876.
- R. N. Kitsis, P. M. Buttrick, E. M. McNally, M. L. Kaplan, and L. A. Leinwand (1991), *Proc Natl Acad Sci USA* 88, 4138–4142.
- C. J. Kuo, P. B. Conley, L. Chen, F. M. Sladek, J. E. Darnell, Jr., and G. R. Crabtree (1992), *Nature* 355, 457–461.
- A. Leask, M. Rosenberg, R. Vassar, and E. Fuchs (1990), *Genes Dev* 4, 1985–1998.
- D. Mischoulon, B. Rana, N. L. R. Bucher, and S. R. Farmer (1992), *Mol Cell Biol* 12, 2553–2560.
- K. T. Montgomery, J. Tardiff, L. M. Reid, and K. S. Krauter (1990), *Mol Cell Biol* 10, 2625–2637.
- L. H. Reid, R. G. Gregg, O. Smithies, and B. H. Koller (1990), *Proc Natl Acad Sci USA* 87, 4299–4303.
- R. W. Tsika, J. J. Bahl, L. A. Leinwand, and E. Morkin (1990), *Proc Natl Acad Sci USA* 87, 379–383.
- L. E. Waspe, C. P. Ordhal, and P. C. Simpson (1990), *J Clin Invest* 85, 1206–1214.
- A. Weydert, P. Barton, A. J. Harris, C. Pinset, and M. Buchingham (1987), *Cell* 49, 121–129.
- J. A. Wolff, R. W. Malone, P. Williams, W. Chong, G. Acsadi, A. Jani, and P. L. Felgner (1990), *Science* 247, 1465–1468.
- K. G. Xanthopoulos, J. Mirkovitch, T. Decker, C. F. Kuo, and J. E. Darnell, Jr. (1989), *Proc Natl Acad Sci USA* 86, 4117–4121.
- K. Zimmerman, E. Legouy, V. Stewart, R. DePinho, and F. W. Alt (1990), *Mol Cell Biol* 10, 2096–2103.