Nucleic Acids Research, 1995, Vol. 23, No. 16 3301–3309

Position independent expression and developmental regulation is directed by the β myosin heavy chain gene's 5′ upstream region in transgenic mice

Stephanie Knotts, Hansjörg Rindt and Jeffrey Robbins*

Department of Pediatrics, Division of Molecular Cardiovascular Biology, Children's Hospital Research Foundation, Cincinnati, OH 45229-3039, USA

Received March 23, 1995; Revised and Accepted July 10, 1995

ABSTRACT

Transgenic mice generated with constructs containing 5.6 kb of the β myosin heavy chain (MyHC) gene's 5′ flanking region linked to the *cat* reporter gene express the transgene at high levels. In all 47 lines analyzed, tissue-specific accumulation of chloramphenicol acetyltransferase was found at levels proportional to the number of integrated transgene copies. Deletion constructs containing only 0.6 kb of 5′ upstream region showed position effects in transgenic mice and did not demonstrate copy number dependence although transgene expression remained muscle-specific. The 5.6 kb 5′ upstream region conferred appropriate developmental control of the transgene to the cardiac compartment and directs copy number dependent and position independent expression. Lines generated with a construct in which three proximal *cis*-acting elements were mutated showed reduced levels of transgene expression, but all maintained their position independence and copy number dependence, suggesting the presence of distinct regulatory mechanisms.

INTRODUCTION

Transgenic mice have been important tools for studying the regulation of gene expression. Regulatory interactions important for tissue-specific gene expression, temporal expression patterns and responsiveness to physiological cues can all be analyzed *in vivo*. In addition, because transgenes are integrated into the chromosome, the effect of chromatin structure on gene expression can be studied. A limitation of transgenic studies is the requirement for analysis of multiple lines due to the fact that, in most cases, the expression pattern of the transgene is at least partially dependent on its integration site. For example, most transgenes require integration into a region of open chromatin in order to direct expression. In addition, *cis*-acting sequences surrounding the integration site may have a dominant effect on the specific spatial and temporal expression pattern of the transgene (1–4).

Although most transgenes show position effects, the flanking sequences of several genes have been shown to direct position independent expression (5–13). This was first demonstrated for the β globin gene and the DNA elements responsible for generating position independence were termed the locus control region (LCR) (9). LCRs are thought to stabilize open chromatin such that the transgene is expressed independent of its site of integration. In addition, many of the *cis*-acting regions responsible for directing position independent expression show an associated ability to direct copy number dependent gene expression (5,7,9,11,12).

Genes which are developmentally regulated may require additional levels of control. The role that an LCR may play in this regulation is not completely understood. In the case of the β globin locus, the LCR interacts sequentially with the promoters of the gene cluster in order of their developmental expression (14,15, reviewed in 16). This suggests that competition for the LCR plays a role in the correct temporal gene expression. In addition, repressors within the promoter region may be required to interrupt the interaction between the promoter and LCR to down-regulate gene expression (17). While chromatin structure is certainly an important mediator of developmental gene regulation, the sequence of events required to initiate the changes in structure are not understood.

The α and β myosin heavy chain (MyHC) genes are found in tandem in the genome (18) and encode isoforms which homo- or heterodimerize to form a component of the thick filament of the sarcomere. An antithetic regulation of these genes occurs in the murine ventricles. The α MyHC gene is constitutively expressed in the atria. The β MyHC gene is expressed in the ventricles during embryonic development. At birth, the β MyHC gene is down-regulated and the α MyHC gene is up-regulated in the ventricles (19). Circulating levels of thyroid hormone may play an important role in this antithetic regulation. The down-regulation of the β MyHC gene at birth and concomitant up-regulation in α MyHC gene expression coincides with a surge in circulating thyroid hormone levels. In the adult, decreased levels of thyroid hormone leads to a re-induction of β MyHC gene expression in the ventricles at the expense of α MyHC transcription (20–23). While thyroid response elements in the α MyHC gene's promoter region seem to play a role in regulating this gene's expression

* To whom correspondence should be addressed

(24–28), it is unclear how the counter-regulation of the β MyHC gene is mediated.

In addition to its expression in the ventricles prior to birth, the β MyHC gene is also expressed in slow (Type I) skeletal muscles in the adult. The gene is therefore regulated at the level of transcription in a manner which allows for its expression in both the cardiac and skeletal muscle compartments at different developmental time points and in response to a variety of physiological cues. We previously undertook a study of β MyHC gene regulation using transgenic mice such that the developmental and hormonal switches in gene expression could be examined. We found that 5.6 kb of the 5′ upstream region of the β MyHC gene was required to direct high levels of tissue-specific and hormone inducible transgene expression (29,30). In addition, some, but not all, of these lines of mice were able to appropriately down-regulate transgene expression in the ventricles at birth.

As part of an ongoing study of the *cis*-acting elements involved in tissue-specific β MyHC gene expression, we mutated three elements in the proximal promoter region both singly and in combination and used these constructs to generate transgenic mice (31). The MCAT element (located at –274 to –284 from the transcriptional start site), and βe3 site (–211 to –189) contain TEF-1 binding sites. The C-rich element (–242 to –237) binds Sp1. Each of these sites was independently required for cell-type specific gene expression in *in vitro* transient transfections (32–34) but did not ablate tissue-specific transgene expression when mutated *in vivo* (31). In the course of our analyses we observed that each of these constructs directed position independent and copy number dependent accumulation of CAT; results that are consistent with the presence of an LCR in the distal 5 kb region. Proper developmental regulation also required the distal 5 kb region, implying the presence of *cis*-acting elements important for cardiac-specific regulation. In addition, in transgenic lines in which all three proximal *cis*-acting elements were mutated, the ability of the promoter to drive high levels of expression was separable from its ability to direct position independent accumulation of CAT.

MATERIALS AND METHODS

Generation and screening of transgenic mice

Mice were generated by pronuclear injection as described previously (35). The β5.6 and β0.6 lines consist of 5.6 and 0.6 kb, respectively, of the 5′ upstream region of the β MyHC gene linked to the chloramphenicol acetyltransferase (*cat*) reporter gene (30). Mutations were made in the MCAT, C-rich or βe3 elements in the proximal promoter using PCR (31). Founder mice were screened by Southern blot analysis as described previously (31). Lines which showed no rearrangements were mated to non-transgenic littermates to establish stable transgenic lines. Offspring were screened by PCR of ear clip DNA using oligonucleotide primers specific for cat (36).

Induction of hypothyroidism

Mice were made hypothyroid by the addition of 0.15% 5-propyl-2-thiouracil (PTU) to their diet (Teklad) for 4 weeks. Hypothyroidism was confirmed by determining the serum thyroid hormone levels using radioimmunoassay (Pentex).

Determination of copy number

Copy numbers were determined by dot blot analysis. Genomic DNA was prepared from tail clips by digestion overnight in lysis buffer (100 mM Tris, pH 8.5, 5 mM EDTA, 0.2% SDS, 200 mM NaCl, 750 μg/ml proteinase K). DNA was spooled and resuspended in TE (10 mM Tris–HCl, pH 7.4, 1 mM EDTA). Genomic DNA was digested with *Eco*RV. Dot blots were performed using a BioRad dot blot apparatus. Two-fold serial dilutions of the digested DNA were applied to a nitrocellulose membrane using 5 μg for the first dot. The DNA was hybridized with a ^{32}P-labeled *cat* fragment (specific activity > 1×10^8). The hybridization signals were quantitated using a PhosphorImager (Molecular Dynamics, ImageQuant V3.3 software). Serial dilutions of a linearized *cat* containing plasmid were applied to each membrane. A standard curve was generated from this calibration plasmid and copy numbers of transgenic lines were determined by comparison.

Preparation of tissue extracts

Tissues were homogenized in 0.5 ml of 250 mM Tris–HCl, pH 7.8 using a Tekmar tissue homogenizer. Fetal tissues were homogenized in 0.2 ml of 250 mM Tris, pH 7.8 by sonication. Extracts were incubated at 65°C for 10 min, centrifuged to remove cellular debris, and used for a Bradford assay (37) to quantitate total protein.

CAT assays

CAT activities were analyzed as described previously (38). Specific protein quantities and incubation times are given in the figure legends. Acetylated products were separated from non-acetylated by thin layer chromatography and the products were visualized by autoradiography.

CAT enzyme-linked immunosorbant assays (ELISAs)

The amount of CAT protein in tissue extracts was quantitated using ELISAs (Boehringer Mannheim). A standard curve was generated using purified CAT and tissue concentrations of CAT (in pg of CAT/μg of total protein) were determined by comparison. For each sample, at least two protein concentrations were measured, both of which fell within the linear portion of the standard curve, and all analyses were performed in duplicate. The variation in independent measurements of CAT protein for a given line was <25%.

RNA analysis

Total ventricular RNA was isolated from the indicated transgenic lines as well as non-transgenic controls using RNAzol (Cinna/Biotecx Laboratories, Inc.) as per the manufacturer's instructions. RNA was spotted in 2-fold serial dilutions on duplicate filters using 6 μg for the first dot. Oligonucleotide probes specific for the α and β cardiac MyHC transcripts (39) were labeled to a specific activity of 2×10^7 d.p.m./pmol using [γ-^{32}P]ATP and T4 polynucleotide kinase. A 320 bp *cat* probe, generated by PCR using oligonucleotides specific for *cat* (36) and identical to that used for genomic dot blots, was labeled to a specific activity >1×10^8 using [α^{32}P]dATP. Hybridizations and washes were performed as described previously (19).

RESULTS

Quantitation of expression and analysis of copy number of transgenic lines

In order to investigate the regulation of the β MyHC gene *in vivo*, a 5.6 kb fragment of the 5′ upstream region was linked to *cat* and used to generate transgenic mice (30). In addition, mutations were made in specific *cis*-acting elements within the proximal promoter region. These constructs are shown schematically in Figure 1A. For each line, the number of integrated transgenes was determined by dot blot analysis. Copy numbers ranged from a low of one copy (confirmed by Southern blot results) to a high of 265 copies (Table 1).

In the adult mouse, β MyHC is expressed in slow (Type I) skeletal fibers although it can be re-induced in the ventricles in response to physiological or hormonal stimuli. We used ELISAs to quantitate the expression of the transgene in the soleus and diaphragm, two skeletal muscles containing a high percentage of slow fibers (Table 1). It has previously been demonstrated that the amount of CAT specific RNA produced correlates well with CAT activity (40–42). To confirm these findings in our transgenic lines, we analyzed *cat* RNA from selected lines of β5.6 wt, β5.6 C-rich, β5.6-βe3, β5.6 triple and β0.6 wt constructs. Representative analyses are shown in Figure 1B. The transcript concentration does not correspond on a 1:1 basis with the protein levels, but, when plotted against protein, a significant correlation (r = 0.993) was obtained. For the β5.6 wt (non-mutated) lines, skeletal muscle expression ranged from 14.35–2307.5 pg of CAT/μg of extract. Similar ranges were observed for all other constructs except the β5.6 triple. Transgene expression in the adult ventricle was also determined for each of the transgenic lines (Table 1). A number of independently generated lines, within each construct, showed ectopic developmental expression (expression in the adult ventricles). The levels of expression in the skeletal muscles and ventricles for each of the constructs varied over three orders of magnitude.

The β MyHC 5′ flanking region directs position independent and copy number dependent expression to skeletal muscles

In order to determine if a correlation between copy number and levels of CAT expression in the skeletal muscles existed, linear regression analysis was performed. Figure 1C shows an example using data for the β5.6 wt construct. There was a highly significant correlation between copy number and skeletal muscle expression with r = 0.988 (p = 0.0001). Similar correlations (p < 0.008 in all cases) were observed for each of the other constructs (Table 2). In addition, radioactive CAT assays performed using extracts from skeletal, smooth, cardiac and non-muscle tissues demonstrated that all lines from each of the constructs (47 lines total) expressed the transgene in a tissue-specific manner independent of the integration site (data not shown). Thus the 5.6 kb promoter fragment directs position independent expression of the transgene to the skeletal muscles in a copy number dependent manner, a characteristic consistent with the presence of an LCR in the upstream region.

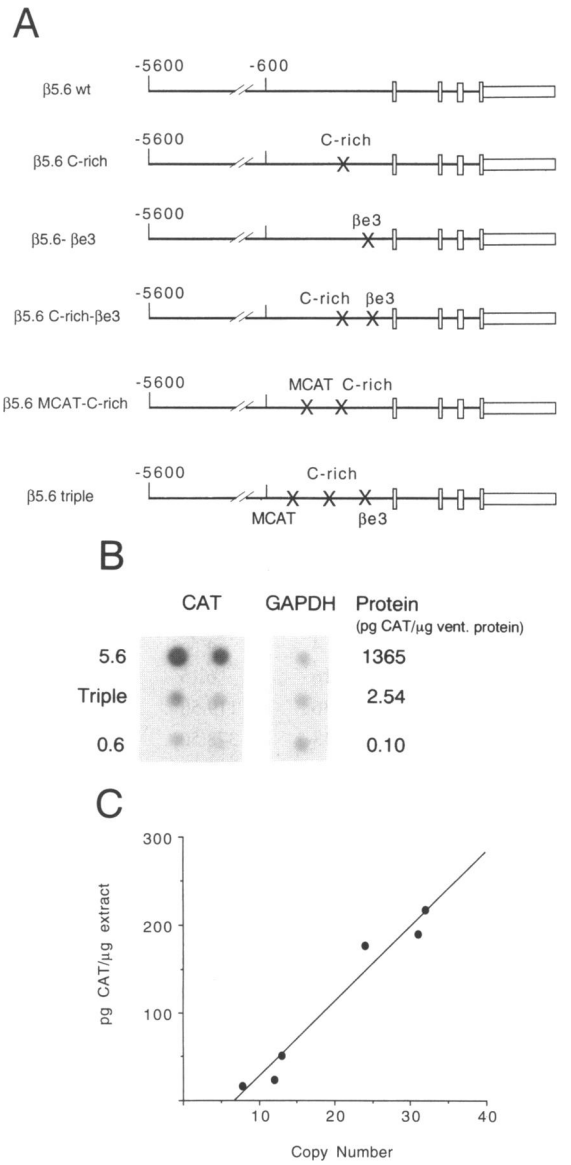

Figure 1. Transgenic constructs and relationship of β5.6 wt skeletal muscle expression and transgene copy number. (**A**) Each construct is shown schematically. Exons 1–4, comprising the complete 5′ untranslated region, are shown as open boxes linked to the cat reporter gene (shaded box). The constructs' name and mutated element(s) are indicated. (**B**) Correspondence of CAT RNA and protein levels. Ventricles from lines carrying either the 5.6 kb promoter fragment (5.6), the 5.6 kb fragment that had been triply mutated (Triple), or the 0.6 kb promoter fragment (0.6) were analyzed for their relative levels of CAT transcript by RNA dot blot hybridization. The probe was a 320 bp *cat* fragment (see Materials and Methods). Six μg were applied to the first well with the second containing a 2-fold dilution. A duplicate blot containing 6 μg of RNA was hybridized to an anti-sense-35mer (GGAACATGTAGAC-CATGTAGTTGAGGTCAATGAAG) corresponding to the glyceraldehyde-3-phosphate dehydrogenase mRNA to assess loading accuracy. The dot intensities (first column) were quantitated using NIH Image 1.57 and the 5.6, triple, and 0.6 lines gave values of 85.3, 37.4 and 20.2 (arbitrary units) respectively. Ventricular protein expression of each line is shown, in pg CAT/μg ventricular protein. Dot intensity plotted against protein on a logarithmic scale gave an r value of 0.993. (**C**) The number of integrated transgene copies was calculated for each line by dot blot analysis. The determinations were made by comparison to a *cat* containing plasmid applied to the same filter (Materials and Methods). For each transgenic line from the β5.6 wt construct, copy number was plotted on the x-axis and CAT expression in the skeletal muscles on the y-axis. Each point represents an independent transgenic line. For scaling purposes line 39 is not represented. The best fit straight line is shown.

Table 1. Copy number and expression characteristics of transgenic lines

Construct	Line Number	Copy Number	pg CAT/μg protein extract Skeletal Muscle[a] Expression[b]	Ventricular Expression[b]
β5.6 wt	34	7	19.95	0.10
	84	12	14.35	0.10
	51	13	50.10	1.33
	46	31	190.55	180.00
	38	32	217.50	70.00
	40	24	177.60	500.00
	39	119	2307.50	2000.00
β5.6 C-rich	12	6	13.70	1.00
	8	12	76.00	1.00
	1	18	1.20	1.00
	24	37	150.00	1.00
	33	64	360.00	77.00
	17	123	655.00	840.00
	34	226	1720.00	1365.00
β5.6-βe3	47	1	2.50	0.65
	61	4	4.55	1.20
	55	16	6.15	0.78
	45	36	119.50	167.00
	59	48	378.50	495.00
	56	80	919.00	1677.00
	62	81	320.00	244.00
	49	87	1538.00	574.00
	68	92	815.00	865.00
β5.6 C-rich-βe3	26	12	0.05	0.13
	35	16	23.60	7.40
	61	17	2.05	1.80
	55	36	9.50	8.60
	44	97	1587.50	1225.00
	31	170	2100.00	3250.00
β5.6 MCAT-C-rich	2	12	5.52	0.09
	13	19	15.40	2.60
	85	23	82.00	3.40
	10	24	229.43	395.00
	91	25	575.85	583.75
	95	28	301.14	255.60
	88	32	320.45	542.50
	86	36	317.00	450.00
	7	86	853.75	570.00
	80	265	1576.25	1685.00
β5.6 triple	58	17	0.27	0.02
	73	18	0.30	0.19
	62	18	4.13	2.54
	42	19	0.10	0.01
	46	22	0.39	0.11
	85	35	50.00	0.72
	67	38	126.00	79.00
	82	80	533.50	146.00

[a]Skeletal muscle expression includes averaged expression data from the soleus and diaphragm

[b]Expression in adult mice (age 6–12 weeks) was quantitated by CAT ELISA. Results are the averages of four to six determinations. Standard deviations in all cases were <25%.

Table 2. Copy number dependence of β5.6 constructs' expression in skeletal muscles

Construct	Number of Lines	Correlation, r, of copy number and skeletal muscle expression levels	p value
β5.6 wt	7	0.988	0.0001
β5.6 C-rich	7	0.987	0.0001
β5.6-βe3	9	0.816	0.0073
β5.6 MCAT-C-rich	10	0.933	0.0001
β5.6 C-rich-βe3	6	0.970	0.0013
β5.6 triple	8	0.951	0.0001

As noted above, the mutation constructs were designed to ablate conserved *cis*-acting elements within the proximal promoter which *in vitro* transient transfections had shown were required for gene activity. Surprisingly, the mutation lines retained their tissue specificity (31). In addition, mutation of these proximal elements did not affect their copy number dependence or position independence (Table 2). This implies that these elements or interaction with these elements are not required to generate position independent and copy number dependent gene expression.

Transgene expression driven by a short 600 bp promoter construct is subject to positional effects and is not copy number dependent

We wished to determine whether a deletion construct containing only 600 bp of the β MyHC gene's promoter retained the properties of copy number dependence and position independence. This region is sufficient to confer tissue-specific transgene expression, albeit at very reduced levels (30). Eleven transgenic lines were generated with the β0.6 wt construct; five expressed the transgene in the slow skeletal muscles of adult mice, whereas in the other six lines no expression in any tissue could be detected. We extended these initial analyses to 33 lines of β0.6 mice in which mutations had been made. To ensure that lines expressing low levels of CAT were not mis-classified as non-expressors, each line was tested under hypothyroid conditions, which results in an up-regulation of both the endogenous β MyHC gene and the transgene (29), and re-assayed prior to classification. Fourteen of these lines expressed the transgene in slow type skeletal muscles. The lines carrying the different mutations were all similar to the β0.6 wt lines in terms of the percentage of lines capable of expression (33–50%). Combining these data (Table 3), 19/44 lines were capable of expression in skeletal muscles while 25/44 showed no expression. These values lie within the range noted by other investigators who have analyzed positional effects (2,3). No correlation of copy number with expression could be detected. Lines that expressed the transgene had copy numbers ranging from 1 to 206 whereas non-expressing lines ranged from 1 to 66. Therefore, this 600 bp region is able to drive transgene expression in a tissue-specific fashion but is subject to position effects and, unlike the 5.6 kb constructs, does not demonstrate copy number dependence.

Position independent and copy number dependent expression in the fetal ventricles

The β MyHC gene is expressed in the developing cardiac tube as early as 8 days post coitum (dpc) (43). Gene expression is restricted to the ventricles during cardiac morphogenesis and β

MyHC transcripts reach a maximum at ~17 dpc (19). Whereas the ventricular expression is down-regulated at birth, expression in the skeletal muscles continues to increase, implying that the gene is differentially regulated in skeletal versus cardiac muscle. Analyses of other genes which are expressed in multiple tissues have shown that at least one transgene, the human keratin 18 gene locus, expresses in a copy number dependent manner in only a subset of its characteristic tissues with some constructs (11). For this reason, we wished to determine if 5.6 kb of the 5′ flanking region of the β MyHC gene was sufficient to direct position independent and copy number dependent transgene expression to the fetal heart in addition to adult skeletal muscles.

Table 3. Expression characteristics of constructs with 600 bp of promoter

Construct	Lines analyzed	Lines showing skeletal muscle expression
β0.6 wt	11	5
β0.6 MCAT	2	1
β0.6 C-rich	8	4
β0.6-βe3	4	2
β0.6 MCAT-C-rich	1	0
β0.6 C-rich-βe3	9	3
β0.6 MCAT-C-rich	3	1
β0.6 triple	6	3
TOTAL	44	19

For each construct, fetuses from at least three independent lines were harvested at 17 dpc and assayed for CAT activity in the heart, liver (a non-expressing tissue) and hindlimb. Representative data for the β5.6 wt construct demonstrate that the lines are able to direct expression to the 17 dpc fetal heart in a tissue-specific manner (Fig. 2A). At this developmental stage, β MyHC is expressed at very low levels in the developing limb muscles. All constructs tested showed similar results, suggesting that 5.6 kb of the 5′ upstream region is sufficient for correct developmental stage-specific expression as well as position independent expression in the heart. To quantitate transgene expression, CAT ELISAs were performed (Fig. 2B). Each line showed a distinct level of expression with high copy number lines directing higher levels of expression. Thus 5.6 kb of promoter of the β MyHC gene appears to direct position independent, copy number dependent expression to 17 dpc hearts.

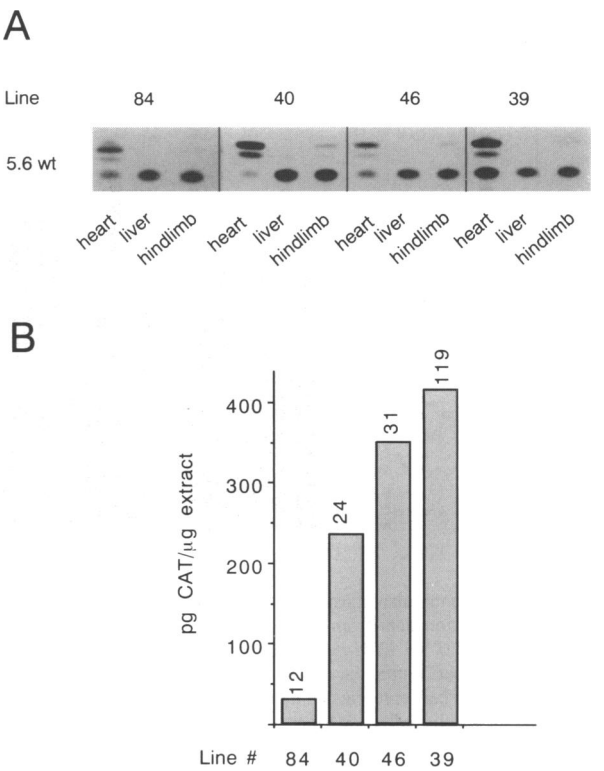

Figure 2. β5.6 wt fetal transgene expression. (**A**) CAT assays performed on tissue extracts from heart, liver and hindlimb of 17 dpc transgenic mice. Four independent lines were tested. Reaction conditions were as follows: line 84, 7.5 μg of protein extract incubated at 37°C for 1 h; line 40, 5 μg incubated 5 min; line 46, 5 μg incubated 5 min; line 39, 2 μg incubated 5 min; (**B**) Quantitation of CAT in 17 dpc transgenic hearts. Hearts were harvested from 17 dpc fetuses. Protein extracts were used in an ELISA to determine the amount of CAT present. For each line the copy number is indicated above the bar.

Transgene expression in the adult heart: correct developmental modulation is related to copy number

In the mouse, the β MyHC gene is down-regulated in the ventricles at the time of birth. However, in some of the β5.6 wt lines the expected down regulation did not occur and the transgene continued to be expressed at high levels in the adult ventricle (Table 1). Approximately 50% of the lines carrying the different mutations also failed to down regulate *cat* expression in the ventricles at birth. Linear regression analyses of copy number and ventricular expression demonstrated that a significant correlation ($p < 0.04$ in all cases) between copy number and expression in the adult heart existed (Table 4). The expression data for the constructs were grouped in order to examine the relationship between correct developmental regulation and copy number (39 lines total). When ventricular expression data was displayed graphically in order of increasing copy number (Fig. 3), a sharp boundary was apparent. Lines with ≤23 copies (15 lines total), all of which expressed the transgene in the ventricles at 17 dpc, appropriately down-regulated expression at birth. For example, β5.6 wt line 84 expressed 35.5 pg of CAT/μg of extract at 17 dpc (Fig. 2B) but down-regulated expression such that in the adult ventricle only 0.10 pg of CAT/μg of tissue extract was present (Table 1). In contrast, lines with ≥24 copies (24 lines total) continued to express the transgene in the adult ventricle at moderate to high levels in a copy number dependent manner. The correlation between the number of transgene copies and the ability to correctly down-regulate transgene expression during development holds, although two exceptions were noted, a line containing 36 copies from the β5.6 C-rich-βe3 construct and a line containing 37 copies from the β5.6 C-rich construct. These exceptions did not appear to be due to the mutations as there were other lines from these constructs which expressed at high levels with high copy numbers. Interestingly, lines containing >23 copies continued to (aberrantly) express the transgene in a copy number dependent manner ($p = 0.0001$). These lines showed a 50-fold range in their levels of expression from 70 pg of CAT/μg of extract to 3250 pg of CAT/μg of extract.

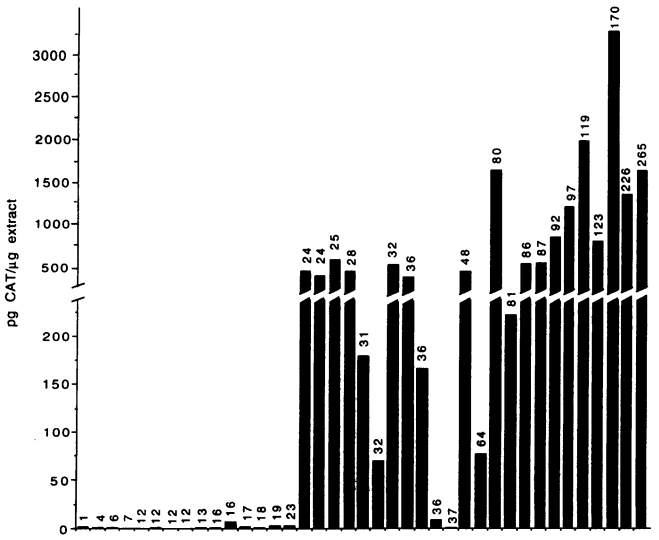

Figure 3. Adult ventricular transgene expression is related to copy number. Ventricular expression data from the β5.6 wt, β5.6 C-rich, β5.6-βe3, β5.6 C-rich-βe3 and β5.6 MCAT-C-rich constructs were pooled and arranged in order of copy number. Ventricular expression is shown, in pg CAT/µg of extract on the y-axis. Each bar represents an independent line. The copy number of a given line is shown above the bar. Lines which have copy numbers >23 express the transgene in the ventricles in a copy number dependent manner ($p = 0.0001$).

Table 4. Copy number dependence of β5.6 constructs' expression in the adult ventricles

Construct	Number of lines	Correlation, r, of copy number and ventricular expression levels	p value
5.6 wt	7	0.967	0.0004
5.6 C-rich	7	0.974	0.0002
5.6 βe3	9	0.704	0.0340
5.6 MCAT-C-rich	10	0.911	0.0002
5.6 C-rich-βe3	6	0.978	0.0007

The endogenous MyHC genes are appropriately regulated in transgenic lines

A surge in circulating thyroid hormone prior to birth is thought to be responsible for transcriptionally activating the α MyHC gene. The mechanism by which the β MyHC gene is concomitantly down-regulated is unknown although it has been postulated that the developmental modulation could occur via the binding of a repressor protein. We hypothesized that the titration of such a repressor could explain the continued transgene expression in the adult ventricles of high copy number lines and that titration would also result in an up-regulation of the endogenous β MyHC gene. To test this, we determined steady state levels of the β MyHC transcript in adult ventricular RNA from six lines of mice: three high expressing lines (>1200 pg/µg ventricular extract), 1 medium expressing line (500 pg/µg ventricular extract) and two lines which appropriately down-regulated transgene expression at birth (<1 pg/µg ventricular extract). The data are presented in Figure 4 and show that, in all samples, the β MyHC transcripts

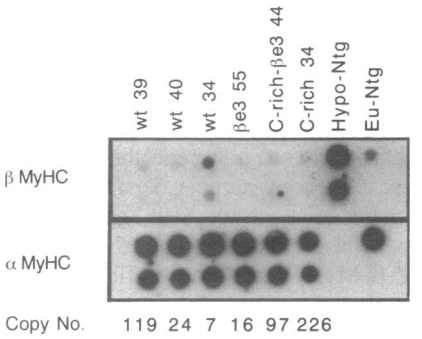

Figure 4. The endogenous MyHC transcripts are unaltered in transgenic lines. RNA was isolated from the ventricles of six transgenic lines, a hypothyroid non-transgenic control and an euthyroid non-transgenic control. Ventricular RNA from each line was used to perform a RNA dot blot. Duplicate filters were prepared using 2-fold serial dilutions (6 µg for the first application) for each sample. One filter was hybridized using an α MyHC specific probe (lower panel) while the second was hybridized using a β MyHC specific probe (upper panel). A single dot containing 6 µg of RNA was applied for the euthyroid non-transgenic control. RNA from the ventricles of a non-transgenic hypothyroid mouse applied to the same filter indicated that the β probe was capable of detecting the transcript. The copy number of each transgenic line is shown below the panels. Hypo-Ntg, hypothyroid non-transgenic ; Eu-Ntg, euthyroid non-transgenic.

were present at normal (very low) levels. In all cases, transcripts from the transgenic animals were indistinguishable from the non-transgenic control and the α and β MyHC transcript complement was unaltered. Therefore, titration of a repressor does not appear to explain the adult ventricular transgene expression seen in these lines.

Copy number dependence is separable from high level expression

The 5.6 kb of 5′ upstream region of the β MyHC gene directed position independent and copy number dependent gene expression to both skeletal and cardiac muscle. Analysis of the 0.6 kb promoter constructs demonstrated that the upstream 5 kb is responsible for this function. In addition, the upstream 5 kb region is required for high levels of expression (~100-fold higher than levels directed by the 0.6 kb promoter) (30). To determine whether these two properties, copy number dependence and high level expression, were distinct we examined copy number and expression levels in triple mutation lines. In adult mice, five of eight triple mutation lines expressed only minimal amounts of CAT (<5 pg/µg of tissue extract) in skeletal muscles, whereas β5.6 wt lines of similar copy numbers expressed at 10- to 1000-fold higher levels (50–190 pg/µg of tissue extract) (Table 1). Three lines of β5.6 triple with higher copy numbers showed increased levels of expression but still remained 4–10-fold reduced as compared to β5.6 wt lines of similar copy numbers. In the fetal heart, a decrease in transgene expression was also seen in the β5.6 triple lines although expression remained tissue-specific (Fig. 5). However, in both the heart and in the skeletal muscles the mutated promoters continued to drive copy number dependent expression (Table 2, skeletal muscle copy number dependence $p = 0.0001$; Fig. 5B). Therefore, mutation of three proximal promoter elements has a significant effect on expression but does not alter copy number dependence.

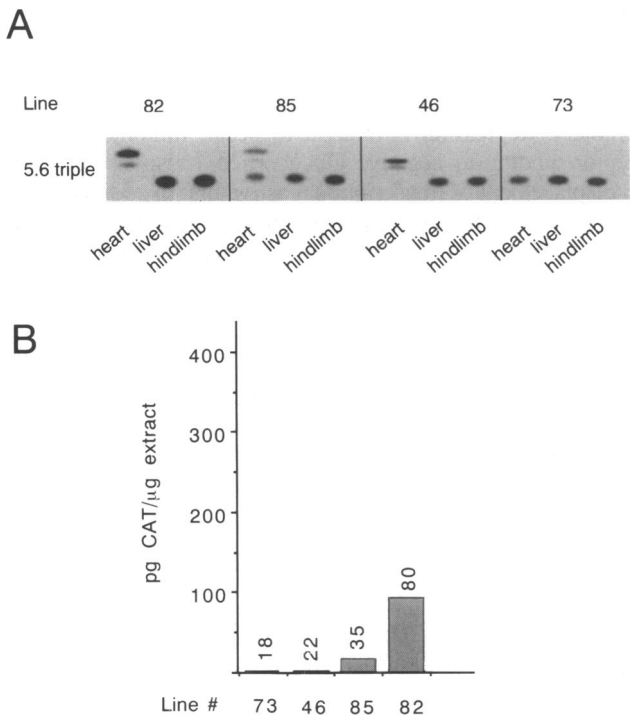

Figure 5. β5.6 triple fetal transgene expression. (**A**) CAT assays were performed on tissue extracts from heart, liver and hindlimbs of 17 dpc transgenic mice. Reaction conditions were: line 82, 7.5 µg incubated 1 h; line 85, 7.5 µg incubated 1 h; line 46, 30 µg incubated 5 h; line 73, 30 µg incubated 5 h. (**B**) Quantitation of CAT in 17 dpc transgenic hearts. Protein extracts were prepared from 17 dpc transgenic hearts and used in an ELISA. The scale is identical to that used for β5.6 wt lines (Fig. 2B) to facilitate comparisons. Copy numbers are shown above the bar for each line.

DISCUSSION

The β MyHC gene is expressed in a tissue- and developmental stage-specific manner. The mechanisms which underlie this tightly controlled expression pattern *in vivo* are not clear and we have used transgenic mice to begin to investigate this regulation. We have analyzed two basic groups of transgenic mice generated with constructs containing either 0.6 or 5.6 kb of the 5′ upstream region of the β MyHC gene linked to *cat*.

A 5 kb fragment of the β MyHC gene's 5′ flanking region directs position independent and copy number dependent expression

Forty-seven lines of mice were generated with the 5.6 kb fragment and all expressed the transgene in skeletal muscles and in the ventricles, suggesting that the *cis*-elements required for muscle specificity are contained within the 5.6 kb fragment and that surrounding chromosomal DNA does not influence *cat* expression, as measured by the accumulation of CAT. All constructs generated with the 5.6 kb fragment directed copy number dependent CAT accumulation, implying that the transcriptional unit is insulated and the locus can function as an independent regulatory unit. In contrast, lines of mice generated using constructs containing only 0.6 kb of promoter region showed position effects. Only 43% of the lines (19/44) showed significant CAT accumulation, presumably due to their integration into open chromatin regions or regions in which genes were being actively transcribed. There was no correlation between copy number and CAT expression levels for the 0.6 kb constructs. Therefore, the distal 5 kb appears to contain the element(s) responsible for position independence and copy number dependence.

We previously demonstrated that the distal 5 kb region of the β MyHC promoter was required to generate high levels of expression (30). Because *cat* is utilized in these transgenic mice, direct comparisons to the endogenous β MyHC gene are not feasible. However, indirect evidence suggests that the 5.6 kb promoter is sufficient to drive levels of expression surpassing that of the endogenous gene. Both the α and β MyHC promoters have been characterized in transgenic mice and their promoter activities quantitated by ELISAs. The β MyHC promoter drives expression levels ranging from 10–2500 pg CAT/µg of total protein (Table 1) whereas the α MyHC promoter drives levels ranging from 600–1300 pg CAT/µg of total protein. The α MyHC promoter has been used to overexpress the ventricular myosin light chain-2, the α_{1B} adrenergic receptors and the β2-adrenergic receptors in the cardiac compartment of transgenic mice (44–46). In each case, the promoter drove levels of expression which surpassed the endogenous gene: the sarcomeric ventricular myosin light chain-2 transcripts accumulated at 4–10-fold that of the endogenous light chain; levels of α_{1B} receptors were 3-fold increased whereas β2-adrenergic receptors were increased 55–195-fold over endogenous levels. Therefore the MyHC promoters appear able to drive high levels of expression in transgenic mice.

Achieving high levels of position independent expression of a transgene is a complex process which most likely requires combinatorial interactions. First, chromatin must be opened such that the transgene is accessible to *trans*-acting factors. Next, the opened chromatin must be stabilized in order to be maintained. It has been postulated that the primary role of the LCR is to foster this stabilized state by entering into cooperative binding interactions which keep the promoter free of histones (reviewed in 47,48). Finally, directing high levels of tissue-specific gene expression often requires the interaction of distant enhancers, with their associated *trans*-acting factors, with the promoter region in order to stabilize the transcriptional apparatus.

The location of an LCR may be marked by DNase I hypersensitive sites. In other cases, such as the human adenosine deaminase gene, the fragment which confers the position independent expression is characterized by Alu-type repeats and is not sensitive to DNase I (5). Both ubiquitous and cell type specific proteins bind in the region of an LCR but no 'LCR-specific binding proteins' have been identified. Although consensus binding sites for both muscle-specific and ubiquitous factors can be found within the β MyHC gene's 5′ upstream region, the elements which confer position independence, copy number dependence, muscle specificity and high level expression await definition. A single Sp1 site is located within the distal 5 kb region. As Sp1 sites can mediate looping, this site may be important for interactions between the distal and proximal promoter regions (49–51). Identification of critical *cis–trans* interactions within the upstream region will require biochemical analyses and given the previous discrepancies between regulatory elements defined *in vitro* and their role *in vivo*, functional analysis

of the upstream elements will require extensive *in vivo* characterization.

Requirements for proper skeletal-, cardiac- and developmental stage-specific control

Previous *in vitro* transient transfections demonstrated the importance of three *cis*-acting elements, an MCAT, C-rich and βe3 site, in β MyHC gene regulation (32–34). These elements are all located within the proximal 600 bp and are therefore contained within the β0.6 constructs. We previously tested the role of these elements in regulating the tissue specificity of the β MyHC gene by mutating each element, alone or in combination, in the context of both the 0.6 and 5.6 kb promoter (31). Mutations in the context of the 0.6 kb promoter had dramatic effects, either causing a loss of skeletal muscle specificity or a loss of transgene expression. However, the same mutations in the context of the 5.6 kb promoter had no effect on either the skeletal muscle specificity or the overall level of expression of the transgene, suggesting that elements in the distal 5 kb region were important for skeletal muscle-specific gene expression and are able to override the proximal *cis*-regulatory elements.

The elements important for regulation of cardiac-specific expression also appear to lie within the distal 5 kb fragment. We previously demonstrated that 0.6 kb of promoter was able to direct expression of the transgene to the cardiac compartment at low levels (30). However, there was no modulation in the levels of expression pre- or post-birth, suggesting that cardiac-specific gene regulation did not occur. In contrast, the developmental down-regulation of expression in the cardiac compartment occurs in the lines containing the 5.6 kb fragment, implying that the elements required for this regulation are located within the distal 5 kb region. In addition, the 0.6 kb promoter is inducible in the skeletal muscles but not in the cardiac compartment in response to decreased levels of circulating thyroid hormone (29). Induction in the cardiac compartment under hypothyroid conditions occurs in lines containing the 5.6 kb fragment, confirming that the element(s) important for cardiac-specific inducibility lie within the distal 5 kb region. Therefore element(s) in the distal 5 kb region appear critical for cardiac-specific gene regulation. As the developmental down-regulation coincides with a surge in the levels of thyroid hormone, this region may contain elements which, directly or indirectly, mediate thyroid hormone responsiveness.

The α and β MyHC genes are located in tandem in the genome, with the β MyHC gene lying upstream. A 4.5 kb intergenic region separates the two genes. Given that these genes are linked and have opposite expression profiles, it has been tempting to speculate that the genes are coordinately regulated although a mechanistic basis for the antithetic expression pattern has not been established. Our data are consistent with the presence of an LCR in the upstream region of the β MyHC gene and this presents one mechanism by which the coordinated regulation could be achieved. In analogy to the β globin locus, in which the genes are organized 5′ to 3′ in order of their developmental expression pattern, developmental stage-specific MyHC gene expression may require the interaction of the LCR with the promoter region of the expressed gene. Alternatively, the elements responsible for the thyroid hormone response may be critical players in the antithetic regulation both during development and in the adult. Other hormone receptors, such as the glucocorticoid receptor, are thought to bind positioned nucleosomes, rearrange chromatin and thereby facilitate transcriptional activation (52,53). The glucocorticoid receptor can also activate or repress transcription of a group of genes by binding to response elements in the genes' promoters. Repression is thought to occur, in some cases, by receptor binding to negative glucocorticoid response elements and displacement of transcriptional activators. In addition, c-Jun and the glucocorticoid receptor regulate gene expression in an antagonistic manner to achieve the proper differentiation and growth during development (54,55). Analogous mechanisms of regulation may be involved in controlling α and β MyHC gene expression. While several thyroid receptor binding sites have been identified in the β MyHC 5′ flanking region by gel mobility shift assays and DNAse I footprinting with bacterially expressed receptors, mutation of these sites did not affect regulation by thyroid hormone in cell culture (56). Although the mechanisms underlying the antithetic regulation of the α and β MyHC genes are unknown, it seems likely that higher order structure plays a role and discrepancies between *in vitro* and *in vivo* data may be explained, in part, by the requirement for physiological chromatin structures.

The lines containing 5.6 kb of the 5′ flanking region showed proper developmental control although this regulation was lost in lines containing >23 copies of the transgene. This suggested that titration of a repressor might play a role in the developmental switch. However, the endogenous α and β MyHC genes continued to be appropriately expressed in these lines. Therefore, it seems unlikely that binding of a titratable repressor within the 5.6 kb fragment effects the down-regulation of β MyHC gene expression in the ventricle and more complex interactions or other mechanisms of regulation must exist. It is possible that high copy numbers affect the integration into chromosomal domains.

Distinct elements in the upstream region control copy number dependence and high level expression

Distinct regulatory regions of the β MyHC 5′ upstream region may be important for mediating position independence versus high level expression and muscle specificity. Although the proximal *cis*-acting elements' mutation did not affect tissue-specific expression or copy number dependence in our β5.6 constructs, simultaneous mutation of all three elements did reduce the overall level of expression. Mutation of these three proximal promoter elements may abolish functional interactions with distal elements important for high levels of expression. In these lines of mice no alteration of the copy number dependence, position independence or tissue-specificity was seen. The data suggest that the generation of high level, tissue-specific, copy number dependent and position independent expression is a complex process and is regulated by distinct elements in the β MyHC promoter. This 5.6 kb fragment, which is not active at low copy numbers but directs copy number dependent expression at higher copy numbers, will be a useful reagent for targeting transgene expression to the adult ventricles over a range of expression levels.

ACKNOWLEDGEMENTS

This work was supported by National Institutes of Health Grants HL22619, HL41496 and HL46826 and by the Marion Merell Dow Foundation.

REFERENCES

1 Al-Shawi, R., Kinnaird, J., Burke, J. and Bishop, J.O. (1990) *Mol. Cell. Biol.* **10**, 1192–1198.
2 Allen, N.D., Cran, D.G., Barton, S.C., Hettle, S., Reik, W. and Surani, M.A. (1988) *Nature* (London) **333**, 852–855.
3 Bonnerot, C., Grimber, G., Briand, P. and Nicolas, J.F. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6331–6335.
4 Palmiter, R.D. and Brinster, R.L. (1986) *Annu. Rev. Genet.* **20**, 465–499.
5 Aronow, B.J., Silbiger, R.N., Dusing, M.R., Stock, J.L., Yager, K.L., Potter, S.S., Hutton, J.J. and Wiginton, D.A. (1992). *Mol. Cell. Biol.* **12**, 4170–4185.
6 Bonifer, C., Vidal, M., Grosveld, F. and Sippel, A.E. (1990). *EMBO J.* **9**, 2843–2848.
7 Carson, S. and Wiles, M.V. (1993). *Nucleic Acids Res.* **21**, 2065–2072.
8 Chamberlain, J.W., Vasavada, H.A., Ganguly, S. and Weissman, S.M. (1991). *Mol. Cell. Biol.* **11**, 3564–3572.
9 Grosveld, F., van Assendelft, G.B., Greaves, D.R. and Kollias, G. (1987). *Cell* **51**, 975–985.
10 Li, Q. and Stamatoyannopoulos, J.A. (1994). *Mol. Cell. Biol.* **14**, 6087–6096.
11 Neznanov, N., Thorey, I.S.,Cecena, G. and Oshima, R.G. (1993). *Mol. Cell. Biol.* **13**, 2214–2223.
12 Schedl, A., Montoliu, L., Kelsey, G. and SchYtz, G. (1993). *Nature* (London) **362**, 258–261.
13 Sharpe, J.A., Wells, D.J., Whitelaw, E., Vyas, P., Higgs, D.R. and Wood, W.G. (1993). *Proc. Natl. Acad. Sci. USA* **90**, 11262–11266.
14 Behringer, R.R., Ryan, T.M., Palmiter, R.D., Brinster, R.M. and Townes, T.M. (1990). *Genes Dev.* **4**, 380–389.
15 Enver, T., Raich, N., Ebens, A.J., Papayannopoulou, T., Constantini, F. and Stamatoyannopoulos, G. (1990). *Nature* (London) **344**, 309–313.
16 Crossley, M. and Orkin, S.H. (1993). *Curr. Opin. Genet. Dev.* **3**, 232–237.
17 Stamatoyannopoulos, G., Josephson, B., Zhang, J.W. and Li, Q. (1993). *Mol. Cell. Biol.* **13**, 7636–7644.
18 Mahdavi, V., Chambers, A.P. and Nadal-Ginard, B. (1984). *Proc. Natl. Acad. Sci. USA* **81**, 2626–2630.
19 Ng, W.A., Grupp, I.L., Subramaniam, A. and Robbins, J. (1991). *Circ. Res.* **68**, 1742–1750.
20 Chizzonite, R.A. and Zak. R. (1984). *J. Biol. Chem.* **259**, 12628–12632.
21 Izumo, S., Nadal-Ginard, B. and Mahdavi, V. (1986). *Science* **231**, 597–600.
22 Lompre, A.M., Nadal-Ginard, B. and Mahdavi, V. (1984). *J. Biol. Chem.* **259**, 6437–6446.
23 Samuel, J., Rappaport, L., Syrovy, I., Wisnewsky, C., Marotte, F., Whalen, R.G. and Schwartz, K. (1986). *Am. J. Physiol.* **250**, H333-H341.
24 Flink, I. and Morkin, E. (1990). *J. Biol. Chem.* **265**, 11233–11237.
25 Gustafson, T A., Markham, B.E., Bahl, J.J. and Morkin, E. (1987). *Proc. Natl. Acad. Sci. USA* **84**, 3122–3126.
26 Mahdavi, V., Koren, G., Michaud, S., Pinset, C. and Izumo, S. (1989). *UCLA Symp. Mol. Cell. Biol.* **93**, 369–379.
27 Subramaniam, A., Gulick, J., Neumann, J., Knotts, S. and Robbins, J. (1993). *J. Biol. Chem.* **268**, 4331–4336.
28 Tsika, R.W., Bahl, J.J., Leinwand, L.A. and Morkin. E. (1990). *Proc. Natl. Acad. Sci. USA* **87**, 379–383.
29 Rindt, H., Knotts, S. and Robbins, J. (1995). *Proc. Natl. Acad. Sci. USA* **92**, 1540–1544.
30 Rindt, H., Gulick, J., Knotts, S., Neumann, J. and Robbins, J. (1993). *J. Biol. Chem.* **268**, 5332–5338.
31 Knotts, S., Rindt, H., Neumann, J. and Robbins, J. (1994). *J. Biol. Chem.* **269**, 31275–31282.
32 Flink, I.L., Edwards, J.G., Bahl, J.J., Liew, C.C., Sole, M. and Morkin, E. (1992). *J. Biol. Chem.* **267**, 9917–9924.
33 Shimizu, N., Prior, G., Umeda, P.K. and Zak, R. (1992). *Nucleic Acids Res.* **20**, 1793–1799.
34 Thompson, W.R., Nadal-Ginard, B. and Mahdavi, V. (1991). *J. Biol. Chem.* **266**, 22678–22688.
35 Hogan, B., Constantini, F. and Lacy, E. (1986) In *Manipulating the Mouse Embryo: A Laboratory Manual.* p. 89–197, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
36 Walter, C.A., Nasr-Schirf, D. and Luna, V.J. (1989). *Biotechniques* **7**, 1065–1070.
37 Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
38 Gulick, J., Subramaniam, A., Neumann, J. and Robbins, J. (1991). *J. Biol. Chem.* **266**, 9180–9185.
39 Robbins, J., Gulick, J., Sánchez, A., Howles, P. and Doetschman, T. (1990) *J. Biol. Chem.* **265**, 11905–11909.
40 Gorman, C.M., Moffat, L.M. and Howard, B.H. (1982). *Mol. Cell. Biol.* **2**, 1044–1051.
41 Gorman, C.M., Merlino, G.T., Willingham, M.C., Pastan, I. and Howard, B.H. (1982). *Proc. Natl. Acad. Sci. USA* **79**, 6777–6781.
42 Stief, A., Winter, D.M., Stratling, W.H., and Sippel, A.E. (1989). *Nature* (London) **341**, 343–345.
43 Lyons, G.E., Schiaffino, S., Sassoon, D., Barton, P. and Buckingham, M. (1990). *J. Cell Biol.* **111**, 2427–2436.
44 Palermo, J.J., Rindt, H. and Robbins, J. (1995). *Ann. NY Acad. Sci.* **752**, 492–505.
45 Milano, C.A., Dolber, P.C., Rockman, H.A., Bond, R.A., Venable, M.E., Allen, L.F. and Lefkowitz, R.J. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 10109–10113.
46 Milano, C.A., Allen, L.F., Rockman, H.A., Dolber, P.C., McMinn, T.R., Chien, K.R., Johnson, T.D., Bond, R.A. and Lefkowitz, R.J. (1994). *Science* **264**, 582–586.
47 Fedor, M. (1992). *Curr. Opin. Cell Biol.* **4**, 436–443.
48 Felsenfeld, G. (1992). *Nature* (London) **355**, 219–224.
49 Li, R., Knight, J.D., Jackson, S.P., Tjian, R. and Botchan, M.R. (1991). *Cell* **65**, 493–505.
50 Mastrangelo, I.A., Courey, A.J., Wall, J.S., Jackson, S.P. and Hough, P.V.C. (1991). *Proc. Natl. Acad. Sci. USA* **88**, 5670–5674.
51 Su, W., Jackson, S., Tjian, R. and Echols, H. (1991). *Genes Dev.* **5**, 820–826.
52 Adom, J., Carr, K.D., Gouilleux, F., Marsaud, V. and Richard-Foy, H. (1991). *J. Steroid Biochem. Molec. Biol.* **40**, 325–332.
53 Rigaud, G., Roux, J., Pictet, R. and Grange, T. (1991). *Cell* **67**, 977–986.
54 Yang-Yen, H.-F., Chambard, J.C., Sun, Y.-L., Smeal, T., Schmidt, T.J., Drouin, J. and Karin, M. (1990). *Cell* **62**, 1205–1215.
55 Zhang, X-K., Dong,J.-M. and Chiu., J.-F. (1991). *J. Biol. Chem.* **266**, 8248–8254.
56 Morkin, E. (1993). *Circulation* **87**, 1451–1460.