Sequence Elements Required for Transcriptional Activity of the Human Myoglobin Promoter in Intact Myocardium

Rhonda Bassel-Duby, Christian M. Grohe, Michael E. Jessen, William J. Parsons, James A. Richardson, Robert Chao, Jayson Grayson, W. Steves Ring, R. Sanders Williams

To define sequence elements required for myoglobin gene transcription in the intact heart, we examined the expression of a reporter gene under the control of a 380-bp upstream segment (−373 to +7) from the human myoglobin gene in transgenic mouse embryos and after gene transfer into left ventricular myocardium of adult rats. This proximal upstream region was sufficient to direct expression of luciferase selectively in cardiac and skeletal muscle of mouse embryos and to recapitulate the pattern of expression of the endogenous mouse myoglobin gene. This same upstream region was transcriptionally active after injection of plasmid DNA into the left ventricular wall of adult rats. Point mutations within two evolutionarily conserved sequence elements—a cytosine-rich (CCAC-box) motif and an A+T–rich (A/T) motif—severely impaired transcription within the intact heart. Nuclear extracts from neonatal cardiomyocytes contain protein factors that bind to each of these elements in a sequence-specific manner. We conclude that combinatorial interactions between the cognate DNA binding factors that recognize these motifs are necessary for transcriptional activity of the myoglobin upstream region in cardiac muscle. (Circulation Research 1993;73:360-366)

KEY WORDS • gene transfer • human myoglobin gene

Myoglobin is a cytosolic hemoprotein that is abundant in adult cardiac myocytes and oxidative skeletal muscle fibers, in which it probably functions to augment delivery of oxygen for mitochondrial respiration during heavy contractile work. The myoglobin gene is transcriptionally regulated during muscle development and in response to physiological signals generated by muscle activity. Previous studies from our laboratory have defined discrete DNA sequence elements within a 2-kb segment of the 5' flanking region of the human myoglobin gene that are required for transcriptional activation during differentiation of skeletal myogenic cells in tissue culture. In transgenic mice, this 2-kb region is transcriptionally active in cardiac and skeletal myocytes exclusively, but specific sequence elements necessary for transcriptional activity in cardiomyocytes have not been characterized previously.

The goal of this project, therefore, was to identify cis-acting elements from the human myoglobin gene that are required for transcriptional activity in cardiac myocytes within the intact heart. We observed that, in transgenic mice, only 380 bp of the proximal 5' flanking region of the human myoglobin gene is required to recapitulate the pattern of expression of the endogenous mouse myoglobin gene. This same region was sufficient to drive expression of a reporter gene after plasmid injection into the myocardium of adult rats. This technique was used further to map transcriptional control elements responsible for cardiac expression of myoglobin at higher resolution. The results indicate that two sequence elements—a cytosine-rich (CCAC-box) motif and an A+T–rich (A/T) motif that are conserved in other mammalian myoglobin genes—are essential for function of the human myoglobin promoter within the myocardium.

Materials and Methods

Plasmids

The simian virus 40 (SV40) early promoter11 was used to drive expression of the chloramphenicol acetyltransferase gene (pSVCAT) and the β-galactosidase gene (pSV40lacZ). The plasmid pGUPPA.8 (general use plasmid with upstream polyadenylation sequence) was constructed by inserting a synthetic oligonucleotide that reproduces the proximal 58 bp of the human hsp70 promoter12 upstream from the firefly luciferase gene,13 carried in Bluescript KS+ (Stratagene Inc, La Jolla, Calif). The pGUPPA.8 construct also includes an SV40 polyadenylation signal located 5' to the promoter region to eliminate background expression of luciferase resulting from cryptic transcriptional initiation sites within the prokaryotic vector sequences. Plasmid pRSV-LUC uses the Rous sarcoma virus (RSV) long terminal repeat14 to drive expression of the firefly luciferase gene. Plasmid pCK9 was constructed by insertion of the

Received September 11, 1992; accepted April 21, 1993.

From the Departments of Internal Medicine (R.B.-D., C.M.G., W.J.P., J.G., R.S.W.), Surgery (M.E.J., R.C., W.S.R.), Pathology (J.A.R.), and Biochemistry (R.S.W.), University of Texas Southwest Medical Center, Dallas.

Drs Bassel-Duby and Grohe contributed equally to the study.

Correspondence to University of Texas Southwest Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75235-8573 (Dr Williams).
The plasmid p380-LUC was constructed by inserting an upstream fragment (−373 to +7) from the human myoglobin gene (Fig 1) 5′ to the firefly luciferase gene carried in Bluescript KS+. This construction, like pGUPPA.8, also includes an SV40 polyadenylation signal 5′ to the site of insertion of the myoglobin sequence. Nucleotide substitutions, directed at disrupting sequence motifs within the CCAC box, A/T motif, and E box sites, were introduced into p380-LUC using oligonucleotide-directed mutagenesis as described. The sequences of the resulting mutant clones (pCCACmut, pEmut1+2, and pA/Tmut) were confirmed by restriction mapping and DNA sequencing. The DNA segment used for transgenic animal experiments contained the myoglobin-luciferase fusion present in p380-LUC but was engineered in the pGEM-7Z vector (Promega Corp, Madison, Wis) to include the SV40 small t intron and an SV40 polyadenylation signal 3′ to the luciferase coding region. Oocytes were injected with linear DNA purified as an Aat II–Nsi I fragment.

Mouse Oocyte Injections and Embryo Analysis

Mice (strain FVB/N-HSD) were housed in the Animal Resources Center of the University of Texas Southwest Medical Center at Dallas, and experimental protocols were approved by the Institutional Review Board for Animal Research. After microinjection of DNA, fertilized oocytes were implanted into pseudopregnant females. Embryos were examined at a gestational age of 18.5 days by Southern analysis (not shown) to identify animals that had integrated the transgene into chromosomal DNA. Tissues of all transgene-positive animals and two transgene-negative littermates were fixed overnight at 4°C in 0.1 M phosphate-buffered saline and 4% paraformaldehyde, pH 7.5, then dehydrated, cleared, paraffin-embedded, sectioned, mounted, and hybridized to radiolabeled probes, as described previously.7 Hybridization probes were labeled during in vitro transcription in the presence of [35S]-UTP. Luciferase mRNA was detected using a 540-nt antisense RNA transcript complementary to firefly luciferase mRNA (+100 to +639). Endogenous mouse myoglobin mRNA was detected with a 503-nt antisense probe complementary to exon 3 of the mouse myoglobin gene,7 and the corresponding sense strand served as a control for nonspecific binding.

Heart Injection Protocol

Rats (Sprague-Dawley) were housed in the Animal Resources Center of the University of Texas Southwestern Medical Center at Dallas, and experimental protocols were approved by the Institutional Review Board for Animal Research. In initial experiments 8-week-old male Sprague-Dawley rats were anesthetized and intubated as described previously.9 In later experiments, rats were anesthetized by methoxyflurane inhalation. After left lateral thoracotomy and exposure of the left ventricle, 100 μL of buffer containing plasmid DNA dissolved in 5% sucrose, 10 mM Tris, pH 7.4, and 1 mM EDTA was injected into the apical portion of the left ventricular wall using a 27-gauge needle. Animals were allowed to recover and were euthanized 4 days later by pentobarbital injection (120 mg/kg IP).

Injections were performed with 100 μg test plasmid DNA carrying the luciferase gene controlled by various promoters. Experiments comparing activity of mutant myoglobin promoter sequences to that of the wild-type sequence present in pMb380-LUC were performed in two ways: (1) with cotransfection of 50 μg pSVCAT as an internal control for variations in the efficiency of DNA uptake into cardiomyocytes and (2) without cotransfection of pSVCAT, to avoid potential confounding effects of competition by the SV40 promoter for factors involved in transcriptional activity of the myoglobin promoter.

The apical third of each heart was resected, rinsed in Dulbecco’s phosphate-buffered saline at 0°C, and weighed. Ventricular myocardium (300 mg) was minced with a scalpel and suspended in 500 μL ice-cold homogenization buffer containing (mM) potassium phosphate, 100 (pH 7.8); magnesium sulfate, 15; and dithiothreitol 1. The tissue was homogenized for 30 seconds (Polytron, Kinematica, Switzerland), and 10 μL octyl-phenoxypolyethoxylate (Sigma Chemical Co, St Louis, Mo) was added. After centrifugation at 3000g for 10 minutes at 4°C, the supernatant was removed and held on ice.

Fig 1. Schematic representation of the human myoglobin gene. The gene is approximately 10 kb in length and includes three exons. The locations of putative transcriptional control elements within the proximal 5′ flanking region (TATA box [−30], two CANNTG [E-box] motifs [−176 to −149], an A+T–rich motif [A/T] [−160 to 169], and the CCAC box [−204 to −226]) relative to the cap site are shown. Evolutionary conservation of these motifs is illustrated by comparison to upstream sequences of the mouse,16 seal,17 and human2 myoglobin genes. Nonconserved bases and base substitutions that generate mutant forms of the myoglobin promoter (CCACmut, A/Tmut, or E1+2mut) are highlighted.
**Enzyme Assays**

To determine luciferase activity in myocardial cell lysates, aliquots containing 50 µL of each supernatant were mixed with 250 µL assay buffer, and light emission was quantitated in a Beckman LS 5000 TA liquid scintillation counter after addition of 100 µL of 1 mM luciferin (Sigma) dissolved in assay buffer. Fifty microliters of each supernatant also was assayed for chloramphenicol acetyltransferase activity, as described previously.\(^{18}\)

**Preparation of Nuclear Proteins From Isolated Cardiomyocytes**

Hearts were excised from 1 to 2-day-old Sprague-Dawley rats. Cardiomyocytes were dispersed with collagenase II (Warthington Biochemical Corp, Freehold, NJ) and centrifuged through a discontinuous Percoll gradient as previously described.\(^{19}\) The myocytes were plated at 1.5×10⁴ cells per 60-mm Falcon dish in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal bovine serum and antibiotic (3 µg/mL gentamicin).

Nuclear extracts were prepared by washing the cell monolayers three times with isotonic buffered saline following by adding to each 60-mm plate of cells 0.4 mL lysis buffer [20 mM HEPES (pH 7.6), 10 mM NaCl, 1.5 mM MgCl₂, 0.2 mM EDTA, 20% glycerol, 0.1% Triton X-100, 10 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, 10 µg/mL leupeptin, 10 µg/mL pepstatin, and 100 µg/mL aprotinin]. The cells were harvested by scraping the plates with a rubber policeman, placing the cellular suspension in a microfuge tube, and centrifuging for 1 minute at 2000 rpm at 4°C. The pellet was resuspended in nuclear extraction buffer (lysis buffer with 500 mM NaCl), rocked for 1 hour at 4°C, and centrifuged at 12,000 rpm for 10 minutes at 4°C. The supernatant was collected and stored at 4°C. Protein concentration was determined by the Bio-Rad protein assay based on the Bradford\(^{20}\) dye-binding procedure (Bio-Rad Laboratories, Richmond, Calif).

**Gel Mobility Shift Assays**

Mobility shift reactions were prepared in a final volume of 12 µL by combining 1 ng of end-labeled double-stranded DNA (=10⁶ cpm) with nuclear extract (around 3 µg) and 1 µg poly (dl-dc) (Boehringer Mannheim Biochemicals, Indianapolis, Ind) in a binding buffer containing 25 mM HEPES (pH 7.6), 100 mM KCl, 5 mM MgCl₂, 5 mM dithiothreitol, 10% glycerol, and 0.2% Nonidet P-40. Unlabeled competitor oligonucleotides (100 ng) were included in some reactions as indicated in Fig. 5. After 20 minutes of incubation at room temperature, the reactions were loaded onto a 4% native polyacrylamide gel equilibrated in 45 mM Tris–45 mM boric acid–1 mM EDTA. After electrophoresis at 10 V/cm for 40 minutes, the gels were dried and exposed to Hyperfilm-ECL (Amersham Corp, Arlington Heights, Ill) for 20 hours. The sequences of oligonucleotide probes and competitors included the following: CCAC (sense), 5'-GATCAGCGAACAACACCCCAAC-CCCCCTGGT-3'; CCAC (antisense), 5'-GATCCGACGGGGTGGTTGTGCTG-3'; CCACmut (sense), 5'-GGTCAGCACACACCCGGTACCTGTG-3'; CCACmut (antisense), 5'-GTGCT-
Rats. In pilot experiments (not shown), the expression of luciferase activity following introduction of pMb380-LUC was comparable to that of pRSV-LUC and pCK9 and consistently greater than that of pGUPPA8. Considerable variability in the expression of luciferase activity was observed with all three of the active promoter constructs, indicating the importance of an internal standard such as pSVCAT that was included in the subsequent analysis of mutations within the myoglobin promoter. Histological sections of myocardium stained for β-galactosidase activity after transfection of pSV40lacZ (not shown) confirmed reports of others that transgenes are expressed in cardiomyocytes but not in other cell types after plasmid injection into the myocardial wall.

**Mutational Analysis of the Myoglobin Promoter Region in the Intact Rat Heart**

We compared the activity of the wild-type myoglobin promoter in pMb380-LUC to that of constructs in which nucleotide substitutions were introduced into the CCAC box (−226 to −205), an A/T element–like binding site (−169 to −160), or both of two CANNTG (E-box) motifs present within this region (−176 to −148). All three of these sequences are conserved in myoglobin genes from other mammalian species (Fig 1). Previous studies demonstrated that the CCAC box and A/T motifs are required for transcriptional activity of the human myoglobin promoter in sol8 cells, a skeletal myogenic cell line. A requirement for these sequence elements for transcription in skeletal muscle also was seen in primary cultures of chick embryo myotubes (R. Bassel-Duby, Z. German, and R.S. Williams, unpublished data). As shown in Fig 3, disruption of the CCAC box or A/T motif severely impaired transcription of the myoglobin promoter in the intact rat heart, whereas mutations that disrupted the E-box motifs had a less profound effect in the cardiac muscle background.

To determine whether results of the mutational analysis were influenced by competition between the SV40 and myoglobin promoters for limiting concentrations of certain transcription factors, some of these experiments were repeated in the absence of cotransfected pSVCAT. In this separate set of experiments, disruption of the CCAC box (n=9) reduced expression of luciferase to less than 5% of that produced by transfection of p380-LUC (P<.01), whereas mutations within both E-box motifs (n=6) had a less striking effect (39% of p380-LUC, P=NS). These results are similar to the set of experiments performed with cotransfection of
pSVCAT that are reported in Fig 3 and suggest that promoter competition had no major influence on conclusions drawn from the mutational analysis.

Nuclear Proteins From Isolated Cardiomyocytes Bind Specifically to the CCAC and A/T Elements

Gel mobility shift assays (Fig 4) indicated that these sequence elements found to be essential for functional activity of the myoglobin promoter within the cardiac wall bind in vitro to nuclear protein factors isolated from cardiomyocytes. Nucleotide substitutions within these sequence elements that disrupt promoter function also disrupt binding of nuclear proteins, as assessed by competitive binding studies.

Discussion

Myoglobin is a member of a set of genes that are expressed in both skeletal and cardiac myocytes but transcriptionally silent in nonmuscle cells. The molecular mechanisms that govern such restricted expression may include cis-acting transcriptional control elements and transcription factors that are common to both lineages of striated muscle or are unique to one and not the other cell lineage. For example, myogenic determination genes from the basic helix-loop-helix family occupy a pivotal role in skeletal myogenesis but cannot be detected in cardiomycocytes. The transcription factors that serve an analogous function during cardiac development remain to be identified.

We have previously studied the regulation of the myoglobin gene in skeletal muscle in some detail. Previous studies of the mechanisms of its expression in cardiac muscle are less extensive, but experiments in transgenic mice demonstrated that a 2-kb upstream fragment (−2038 to +7) of the human myoglobin gene is sufficient to direct expression of a reporter gene selectively in skeletal and cardiac muscle of developing mouse embryos. This implies that, unlike certain other genes, muscle-specific enhancer regions that control expression of myoglobin in cardiomycocytes reside within the same genomic segment as those that govern expression in skeletal muscle. The present study was designed to identify at higher resolution cis-acting transcriptional control elements required for transcription of the myoglobin gene in cardiomycocytes. For this purpose, we combined additional studies in transgenic animals with the recently developed technique of direct plasmid injection into the myocardium.

The data indicate that a limited segment (−373 to +7) of the upstream region of the human myoglobin gene is sufficient to direct transcription of a reporter gene in the heart. This conclusion is supported both by
analysis of transgenic mice and of rat myocardium following somatic cell gene transfer. Analysis of base substitution mutants indicates that cardiac transcription requires at least two sequence motifs within this region: a CCAC box, and an A/T motif. Consensus E-box motifs within the myoglobin promoter appear to have lesser importance in cardiac muscle, as we have observed previously in skeletal myocytes.6

The CCAC-box sequence is conserved in the promoter region of myoglobin genes from other mammalian species (Fig 1), and similar sequences are present within transcriptional control regions of other genes that are expressed exclusively in skeletal and cardiac muscle.21 Results from the present study indicate that cardiomyocytes express nuclear proteins that bind to the CCAC element in a sequence-specific manner. The identity of these cardiomyocyte proteins remains to be determined definitively but may be related to CCAC-box binding proteins we have characterized from skeletal myocytes in other studies. These include a 40-kD protein purified from so8 myotubes6 and a less abundant 89-kD protein identified by screening bacterial expression libraries with a radiolabeled CCAC sequence as probe. Antibodies raised against this latter protein cross-react with a protein of similar size in Western blots of nuclear proteins from isolated cardiomyocytes (R. Bassel-Duby, M. Hernandez, and R.S. Williams, unpublished data).

Cardiomyocytes also express nuclear protein factors that recognize the A/T motif from the myoglobin upstream region, as shown in the present study. In other investigations (not shown), gel mobility shift assays indicate that this region is capable of binding MEF225,26 but at a lower affinity than a consensus MEF2 element from the mouse creatine kinase gene. This region also binds specifically to other factors present in muscle extracts that can be distinguished from MEF2 by competitive binding assays (J. Grayson, R. Bassel-Duby, and R.S. Williams, unpublished data). Determination of the identity of the factor(s) that recognizes this functionally important control region in cardiomyocytes also is a major goal of ongoing investigations.

E-box motifs are evolutionarily conserved within myoglobin genes of several species. We suspect that binding of basic helix-loop-helix proteins to these sites may have a role in transcriptional activation of the myoglobin gene during skeletal or cardiac muscle development, but this function appears to be largely redundant to that of other factors under the conditions used for analysis in the present study and previous studies6,7 of the myoglobin promoter.

We conclude from the present study that combinatorial interactions between the cognate DNA binding factors that recognize CCAC-box and A/T motifs within the human myoglobin upstream region are necessary for transcriptional activity in cardiac muscle. Further characterization of these factors and elucidation of the nature of this functional interaction in molecular terms should increase our understanding of the control of gene expression in cardiac myocytes and of the developmental biology of the heart.

Acknowledgments

This study was supported by National Institutes of Health Grants HL-35639 and AR-40849 (R.S.W.) and by the Perot Family Foundation. C.M.G. is a recipient of a postdoctoral fellowship from the Deutsche Forschungsgemeinschaft. W.J.P. is the recipient of a Clinician-Scientist Award from the American Heart Association and Genentech. We thank Mark Gonzalez, Linda Benecke, John M. Shelton, and Angela Smiddy for excellent technical assistance. In addition, we thank Zohre German for culturing cardiomyocytes from neonatal rats. Transgenic mice were generated in the Transgenic Animal Facility of the University of Texas Southwestern Medical Center, Dallas, under the direction of Randall W. Moreadith and with the assistance of Kathy Graves. Sharon Youngblood performed Southern blots and made important contributions to in situ hybridization experiments.

References


Sequence elements required for transcriptional activity of the human myoglobin promoter in intact myocardium.
R Bassel-Duby, C M Grohe, M E Jessen, W J Parsons, J A Richardson, R Chao, J Grayson, W S Ring and R S Williams

Circ Res. 1993;73:360-366
doi: 10.1161/01.RES.73.2.360

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circres.ahajournals.org/content/73/2/360