Focus on Molecular Biology

Genetic Mechanisms That Determine Oxidative Capacity of Striated Muscles
Control of Gene Transcription

R. Sanders Williams, MD

The conceptual basis for understanding regulation of gene transcription in mammalian cells has advanced dramatically in recent years, and research in this field has spawned powerful new experimental techniques. This review will summarize some of these advances and illustrate their application to problems of special interest to cardiovascular investigators as well as their increasing relevance to cardiovascular medicine.

Why Study Transcription?

Because the molecular basis of cell diversity within multicellular organisms results primarily from selective transcriptional activation of certain genes and inactivation of others, the general biological importance of efforts to understand transcriptional control in biochemical terms is self-evident. However, clinicians and investigators without expertise in molecular biology may be less aware of the increasing relevance of an understanding of transcriptional control mechanisms for certain areas of clinical medicine. For example, the protein products of several viral oncogenes and cellular proto-oncogenes have proven to be transcription factors. The receptors for steroid and thyroid hormones are themselves transcription factors that are activated by ligand binding.

The need for fluency in understanding control of gene transcription among cardiovascular investigators and clinicians is no less immediate than for oncologists and endocrinologists. Efforts to delineate pathophysiological mechanisms of congenital heart disease, cardiac hypertrophy and failure, angiogenesis, vascular remodeling, arrhythmogenesis, and atherosclerosis frequently focus on transcriptional regulation of certain genes. Furthermore, delineation of transcriptional control mechanisms may permit the development of novel pharmacological therapies designed explicitly to modulate expression of specific genes. Many drugs already in clinical use have the potential to alter transcription, in some cases in a selective manner, and this potential may be relevant to their therapeutic efficacy or toxicity. Finally, the success of gene therapy for cardiovascular diseases will depend not only on novel delivery systems for introduction of foreign genes into patients but also on regulation of their transcription to maximize therapeutic impact.

Regulation of Gene Transcription
General Concepts

Many of the steps in the process by which the genetic information within DNA is used to produce specific proteins are subject to biological regulation. In addition to gene transcription, these regulatable steps include post-transcriptional processing or stability of RNA and translation or post-translational modifications of proteins. However, initiation of gene transcription is the dominant site for control of synthesis of many proteins. Genes encoding ribosomal RNA subunits and transfer RNAs are transcribed by different RNA polymerases and are regulated by somewhat different control mechanisms than those that act on the enzyme [RNA polymerase (Pol) II] that transcribes protein-coding genes. Only this latter class of genes transcribed by Pol II will be considered in this review.

At any given time, only a portion of the genes within a cell are actively engaged in transcription of RNA; others are silent and do not contribute to the specific phenotype of the cell. Some genes are transcribed at an almost constant rate in all cells and are termed "housekeeping genes." Other genes encode proteins that determine the specialized morphology and function of specific types of cells (e.g., neurons, myocytes) and therefore must be subject to stringent controls arising during cell differentiation. Finally, some genes are transcribed in one or more different types of cells, but the rate of their transcription is modulated over a broad range and depends on specific signals received by the cell from the extracellular environment.
A. Cis-acting transcriptional control elements

Transcriptional activity can be modified by chromatin structure and chemical modifications of nucleic acids.6-8 Methylation of nucleotide bases, interactions with cytoskeletal structures (nuclear scaffold), or covalent modifications of histones may act on individual genes or groups of genes to preclude their transcription in certain cell backgrounds. However, the absence of these inhibitory processes usually is not sufficient to induce transcription. When the chromatin structure permits transcription, whether the gene is actually transcribed usually depends on the interaction of diffusible protein factors and nucleotide sequences adjacent to the gene.9-11

The nucleotide sequences that serve this control function are called “cis-acting regulatory elements” because they are present within the same DNA molecule as the gene itself. The proteins that interact with these regulatory elements are termed “trans-acting regulatory factors” because they are free to diffuse through the nucleus and exert their effects on any DNA molecule; they are not confined to the DNA molecule in which they are encoded. This concept is illustrated schematically in Figure 1.

cis-Acting Regulatory Elements

Transcriptional control elements are short sequences of DNA (five to 20 nucleotide base pairs) usually located adjacent to the 5' end of the protein-coding region of a gene. These elements are quite small relative to the size of the entire gene, which typically occupies from several thousand to more than 1 million base pairs. Many genes contain multiple cis-acting regulatory elements, sometimes with extraordinary complexity.

It is useful to divide transcriptional control elements into several functional classes. Core promoter elements, as typified by the TATA box (common form of core promoter element) family,12 are located within a strictly delineated space relative to the transcriptional start site and serve to define the position within the DNA at which transcription is initiated. Although core promoter elements alone are sufficient for basal levels of transcription to occur, in most genes transcription is regulated by members of a second class of control elements. The terminology used to describe members of this second category of regulatory element is still evolving, but they may be grouped together under the term “inducible promoter elements” because their function can be induced by a variety of effector mechanisms and regulatory stimuli.

Like core promoter elements, inducible promoter elements comprise short DNA sequences and are often found adjacent to the transcribed portion of genes within a few hundred base pairs 5' to the transcriptional start site. Control sequences in this location are also termed “upstream regulatory elements” or “upstream activator sequences.” However, unlike core promoter elements, inducible promoter elements do not define the start site of transcription,
and their ability to modulate transcription often is not dependent on a strict spatial relation to the site at which RNA synthesis is initiated. Indeed, inducible promoter elements may reside 3' (downstream) from the core promoter or within the transcribed portion of the gene, within introns, or at vast distances (>10,000 base pairs) from the core promoter. Inducible promoter elements that are active irrespective of their position or location relative to the start site of transcription are called "enhancers."

Inducible promoter elements may be subdivided further into several categories. Certain inducible elements are constitutive in that they function in a manner that is relatively independent of the cell background in which the enhancer is found. Others are cell specific or stage specific in that they function in only certain types of cells or at specific stages of development of a cell lineage. Other elements are responsive only on stimulation by specific biochemical signals (e.g., hormone response elements, metal response elements) or physical stresses (e.g., heat shock elements). Inducible elements that suppress rather than increase the rate of transcription are termed "silencers."

The boundary between core promoter elements and inducible promoter elements is not always distinct. Although the functional distinctions outlined in the preceding paragraphs are useful for heuristic purposes, they are not universally applicable and exceptions exist. For example, in addition to determining the start site for transcription, core promoter elements sometimes mediate regulatory effects of certain biological stimuli.

A variety of hypotheses have been advanced to account for the functional effects of cis-acting regulatory elements on transcription. Certain DNA sequences can form three-dimensional structures that may alter the ability of RNA Pol II to bind DNA and initiate transcription. However, the predominant mechanism by which cis-acting regulatory elements modulate transcription is by providing binding sites for nuclear proteins that act in trans and in a combined manner to initiate transcription.

**Trans-Acting Regulatory Factors**

Core promoter elements bind to discrete nuclear proteins that serve to position the RNA polymerase and several other accessory proteins (that do not bind DNA) on the DNA strand to be transcribed. The multicomponent aggregate that results is called an "active transcriptional complex." For example, the TATA box binds a protein termed "TFIID," which in turn interacts with Pol II and other general transcription factors.

Enhancers and other inducible promoter elements also bind nuclear proteins. In some examples, the enhancer-protein interaction is relatively straightforward. The glucocorticoid receptor, in the absence of hormone, is complexed to another protein that prevents it from binding DNA. On binding glucocorticoids, the receptor becomes localized to the nucleus, where it binds to a hormone-responsive enhancer called the "glucocorticoid response element" (GRE). The binding of receptor protein to a single copy of the GRE is sufficient to activate transcription.

In other examples, the enhancer-protein interaction is quite complex. For example, the 72-base pair repeat region of the SV40 virus, a potent enhancer in most vertebrate cells, comprises multiple protein-binding sites (enhansons) that bind several different proteins. Some enhanson sequences are devoid of enhancing function when studied as single units but are active when present in multiple copies or in combination with different enhansons. Thus, the properties of an enhancer often are dependent on the combined effect of highly complex interactions between multiple transcription factors and their cognate binding sites. Because of the presence of multiple types of transcriptional control elements within most mammalian genes, a relatively small number of different transcription factors have the potential to generate an enormous number of combinations and therefore a complex repertoire of regulatory possibilities.

How do DNA-binding proteins activate transcription, particularly when the enhancer is located hundreds or thousands of base pairs away from the transcriptional start site? Although several hypotheses have been advanced, the most widely accepted model, based on several lines of evidence, is that enhancer-binding proteins activate transcription by direct interaction with RNA Pol II or other protein components (e.g., TFIID) of the complex that forms on the core promoter. Such protein-protein interactions, in a manner that remains to be elucidated, facilitate or stabilize assembly of the active transcriptional complex and increase the rate at which transcription is initiated. In the process, the intervening DNA sequences are looped out.

A variety of proteins that bind to DNA sequences within transcriptional control elements have been purified and cloned. Although this work is too extensive to review in detail, a few common principles are noteworthy. First, the same protein may recognize apparently diverse DNA sequence motifs. For example, the heme-dependent regulatory factor HAP1 of yeast binds specifically to elements within two different genes that have no apparent sequence similarity.

Second, the same DNA sequence motif may bind more than one protein. An example of this principle is the family of proteins that bind a regulatory element known as the CCAAT box. In some genes, this element appears to function as a core promoter element, but in other cases, it is inducible by discrete biochemical signals in a cell-specific manner. Thus, the functional properties of the CCAAT element are not uniform but instead depend on the potential cognate binding proteins that are present. Other evidence, including some of our own findings, suggest that the TATA box also may bind multiple forms of TFIID.

Third, transcriptional regulatory proteins frequently form complexes with other proteins. In several exam-
ples, binding of the transcription factor to a second protein prevents the transcription factor from binding DNA. NF-κB, a protein involved in specific activation of certain genes during lymphoid differentiation, is present in nonlymphoid cells but is inactive because it is complexed to an inhibitory protein. Other transcription factors do not bind DNA directly but instead activate transcription only as a component of a dimeric protein complex. For example, the c-fos protooncogene product is itself devoid of DNA-binding activity, and its transcriptional activation function is realized only when it is complexed to the product of the c-jun protooncogene, forming a transcription factor termed “AP-1.”

Finally, transcriptional activator proteins often can be separated into discrete functional domains. By experimental mutagenesis of the protein, regions responsible for DNA binding can be distinguished from regions required for transcriptional activation, recognition of signals (i.e., hormones or second messengers), complex formation with other proteins, or nuclear localization. A variety of structural motifs are frequently found in DNA-binding proteins and have been given colorful names: zinc finger, copper fist, homeobox, helix-turn-helix, leucine zipper, and acid noodle. By shuffling domains from one transcription factor to another, it is possible to create chimeric proteins with hybrid characteristics.

Domain analysis of transcription factors has produced several important insights but has also raised many new questions. For example, no simple relation between protein structure and DNA binding is evident. The structural relations that determine how proteins distinguish between individual nucleotide sequences to bind with such exquisite specificity remain to be elucidated. Domain analysis also makes it clear that DNA binding is not sufficient to activate transcription. Regulatory elements within the DNA serve to localize transcription factors so they can present their transcriptional activation domains to RNA Pol II or other components of the transcriptional complex, but stimulation of transcription is a separable function from DNA binding. The structural details of the protein-protein contacts that result in activated transcription remain to be defined.

Methods for Study of Transcriptional Regulation

The preceding concepts and models are based on experimental results derived from application of powerful experimental methods, most of which have been developed in the past few years. Informed appraisal of scientific literature concerning transcriptional regulation requires a basic understanding of terminology and operating principles of some of these procedures. Laboratory techniques applicable to the study of transcriptional regulation include a variety of methods for quantifying transcriptional activity after experimental manipulation of DNA sequences thought to contain regulatory elements and several different approaches designed to identify and characterize transcription factors. These techniques are complementary and usually used in concert for study of the transcriptional regulation of any given gene or family of genes.

Functional Assays of Transcriptional Control Elements

Most studies of transcriptional regulation begin with the approach illustrated in Figure 2. Segments of DNA adjacent to the native transcriptional start site of the gene of interest are cloned into plasmid vectors that contain the protein-coding region of a different gene. Bacterial plasmids genetically engineered in this manner are called “constructs.” This protein-coding gene is termed the “reporter,” or indicator, gene. Reporter genes encode proteins that are not present in the type of cells in which transcriptional control will be studied and that can be detected and quantified by convenient and sensitive biochemical or histochemical assays. The bacterial chloramphenicol acetyltransferase28 or β-galactosidase genes29 and the firefly luciferase gene30 are often used for this purpose in mammalian or avian cells. As a confirmatory (or alternative) approach, messenger RNA (mRNA) transcribed from the foreign gene is analyzed directly.

Constructs containing a reporter gene and putative transcriptional control regions are purified and introduced into cells in culture. Cells freshly isolated from laboratory animals are called “primary cultures,” whereas cells propagated during many generations in culture as clonal derivatives are called “established cell lines.” Either type of culture can be suitable for this purpose. The introduction of foreign DNA into cells in culture is called “transfection” and can be accomplished by several methods.

The most formidable step that must be overcome for successful transfection involves transport of the construct across the cell membrane. Once inside the cell, foreign DNA may be degraded, but it also may be transported to the nucleus, where it is accessible to transcription factors. When coprecipitated with calcium phosphate under certain exacting conditions (e.g., pH, DNA purity), foreign DNA is taken up from the extracellular environment by many types of cells.31 Alternatively, the foreign DNA for transfection can be complexed with lipids that fuse with cell membranes and deliver the DNA into the interior of the cell (lipofection).32 The integrity of the cell membrane can be temporarily disrupted by an electrical field (electroporation)33 so DNA from the surrounding medium can enter the cell. In addition, DNA can be coated on the surface of microparticles, which are propelled across the cell membrane by rapid gas expansion (biolistics).34 Finally, constructs can be prepared so the sequences of interest are packaged in infectious viral particles, usually in retroviruses.35 Because the viral particles recognize specific cell surface receptors, this mode of transfection is particularly efficient.

After transfection, reporter gene activity can be measured without further manipulations of the cells.
Functional assays of transcriptional control elements

1. Prepare construct with putative control elements linked to reporter gene

2. Introduce construct into cells

3. Vary conditions under which cells are maintained

4. Assay reporter gene product (mRNA or protein)

Figure 2. Functional assay of transcriptional control elements by transfection of mammalian cells in culture. Recombinant plasmids (constructs) are prepared by fusion of putative transcriptional control elements to protein-coding region of a reporter gene foreign to host cell. These constructs are introduced into cells by one of several methods (see text), and cells are subjected to conditions expected to activate putative control elements. Transcription of foreign gene is quantified by measurement of its messenger RNA (mRNA) or protein product.

(transient expression assays) or after additional steps designed to select cells that have incorporated the foreign DNA into their chromosomes so the foreign sequences are passed faithfully to daughter cells after cell division (stable transformation assays).

The basic logic of these assays is that the reporter gene will be expressed if the construct contains control elements capable of driving transcription in the transfected cell. If no expression of the reporter gene is observed, the experiment is repeated with different segments of the gene of interest until functional regulatory elements are located. Once a region that supports transcription is identified, deletions and mutations are created and analyzed to localize their important control elements more precisely and to distinguish their functional properties.

There are many variations on this theme. Cell-specific control elements are defined by their ability to drive transcription after transfection into one type of cell but not another. Likewise, regulatory elements that respond to hormonal signals (e.g., glucocorticoid, thyroid response elements), chemical signals (e.g., cyclic AMP, heavy metals), or physical stresses (e.g., heat shock, hypoxia) are defined by their ability to increase transcription specifically when the appropriate stimulus is applied. The ability of a specific protein to activate a specific regulatory element can be analyzed by simultaneous transfection (cotransfection) of two constructs—one bearing a target control element fused to a reporter gene and the other bearing an effector gene encoding a putative transcription factor driven by a constitutive promoter.

Transcriptional control elements also can be defined in cell-free systems. DNA constructs can serve as the template for transcription in the presence of purified or partially purified RNA Pol II, accessory general transcription factors, and factors that recognize specific regulatory elements. Such in vitro transcription assays can complement studies performed in living cells and, in particular, can elucidate the detailed biochemistry of transcriptional control.

In addition, as the technology required to generate germline transgenic animals has become available on a widespread basis and as procedures for genetic transformation of somatic cells in living animals are developed and improved, it has become increasingly feasible to study transcriptional control elements in intact laboratory animals. The opportunity to study transcriptional control elements in the context of tissues of living animals, rather than in the highly artificial context of cell culture, is particularly attractive to cardiovascular investigators for several reasons. First, established cell lines that differentiate into cardiac myocytes are not available. Second, cultured cell systems are at best incomplete and, in other cases, are totally inadequate as models for many of the physiologically and clinically relevant stimuli that modulate gene expression in the heart and vasculature. Investigations of transcriptional control mechanisms in transgenic animals will provide powerful new experimental strategies to increase our understanding of mechanisms that control cardiac growth, angiogenesis, and cardiovascular remodeling in response to
physiological stresses; cardiovascular physiologists will be called on to analyze the physiological effects of forced expression of selected proteins in myocardial or vascular tissues.

Techniques to Identify Transcription Factors

Once a region of DNA containing a transcriptional regulatory element has been defined by functional assays, several methods known generically as "footprinting" can help to identify the specific nucleotides that constitute a protein-binding site. Investigators then exploit knowledge of the protein-binding site to identify and characterize the binding proteins. In gel retention assays, DNA fragments or synthetic oligonucleotides are radioactively labeled and incubated with nuclear proteins extracted from cells. If the nuclear extract contains proteins that bind to the labeled DNA, DNA-protein complexes are formed that are retained within the gel when placed in an electrical field under non-denaturing conditions and migrate at a slower rate than unbound DNA (Figure 3).

With appropriate controls to establish specificity, gel retention assays can be used to assess the distribution of DNA-binding proteins among different types of cells, to monitor the success of purification protocols, to establish cofactor requirements and to assess activation of DNA binding by hormonal or physical signals.

The binding of proteins to specific nucleotide sequences is reminiscent of the interaction of a drug or hormone with its receptor or of the recognition of an antigen by a specific antibody. The specificity of the interaction provides an opportunity to select the binding protein from a complex mixture of other proteins through a procedure known as "DNA affinity chromatography." Synthetic oligonucleotides containing multiple copies of a given protein-binding site are linked covalently to an insoluble support resin. Under conditions of ionic strength and pH that favor specific DNA-protein interactions, the protein to be purified binds to the resin while other proteins pass through. The protein of interest can then be recovered by washing the column at higher ionic strength. Proteins that bind specific DNA sequences with high affinity can be purified more than 100-fold in a single step by this procedure.

Transcription factors are rare proteins and are often present in only a few copies per cell. Nevertheless, it has been possible to purify several transcription factors in sufficient quantities to obtain partial amino acid sequences and to subsequently clone complementary DNA and genomic DNA encoding these proteins. In addition, several methods have been developed that potentially permit cloning of complementary DNA encoding DNA-binding proteins without biochemical purification.

Studies of Transcriptional Control Applied to Specific Questions in Cardiovascular Biology: Regulation of Genes Encoding Proteins of Oxidative Metabolism

Physiological and Clinical Relevance

All cells of vertebrate organisms require oxygen to survive and to carry out specialized cellular activities. However, cells differ markedly in their requirements and capacity for generation of ATP by oxidative phosphorylation. At one extreme of this spectrum, cardiomyocytes face enormous energetic demands during sustained, repetitive contractile work. Near the other extreme, type IIb skeletal muscle fibers are adapted for short bursts of contractile activity, which can be fueled by anaerobic glycolysis, and they fulfill their physiological function with minimal requirements for oxidative metabolism.

Notably, the capacity of muscle cells to metabolize oxygen is not fixed during terminal differentiation in the embryo but is responsive to changes in physiological demand in adult animals and humans. Exercise conditioning augments the oxidative capacity of skeletal muscles and enhances the ability to sustain work without fatigue. Conversely, physical inactivity, neuroendocrine abnormalities, and possibly other factors reduce the oxidative capacity of skeletal muscles and contribute to the exercise intolerance manifested by patients with chronic congestive heart failure. The remarkable adaptive capacity of skeletal muscle in this regard has recently been exploited in imaginative approaches to therapy of congestive
heart failure in which skeletal muscles are transplanted into the chest to power ventricular assist devices or artificial ventricles. The success of this approach depends on conditioning the skeletal muscle by chronic nerve stimulation to augment its oxidative capacity and resistance to fatigue.

In the myocardium, certain clinical disorders are attributable, at least in part, to abnormalities in the capacity of cardiomyocytes to conduct oxidative metabolism. Such disorders may arise as a result of inherited gene defects, as a toxic effect of drug administration or nutritional deficiencies, or as a failure of normal adaptive responses to cardiovascular stresses imposed by valvular disease, hypertension, or loss of functional myocardium.

How do striated myocytes determine their capacity to produce ATP by oxidative metabolism during development of specialized muscle subtypes and as an adaptive response to changing physiological demands? For several years, my laboratory has regarded this question as one of fundamental importance for cardiovascular physiology. Our efforts to address this problem have led us into several rapidly advancing areas of modern molecular biology, including mitochondrial genetics, intracellular signaling, growth factors, and the development of novel gene transfer technologies. However, our research has increasingly included efforts to elucidate mechanisms by which transcription of individual genes encoding proteins of oxidative metabolism is regulated. The importance of transcriptional regulation, not only for our specialized interests but also for many areas of cardiovascular research, makes this area of our research appropriate for review in this series.

Genes Encoding Proteins of Oxidative Metabolism

The model of transcriptional regulation based on the interactions of cis-acting regulatory elements with trans-acting factors to form an active transcriptional complex has provided a conceptual framework for our efforts to understand genetic mechanisms that regulate the oxidative capacity of cardiac and skeletal muscles, at least concerning nuclear genes. A small but essential subset of genes that encode proteins of oxidative metabolism are located within mitochondrial DNA. Although these mitochondrial genes have also been an important focus of attention in our laboratory, they appear to be regulated by fundamentally different mechanisms and will not be discussed further in this review.

Our working hypothesis is that nuclear genes of this class respond to biochemical signals generated as a component of programs of muscle differentiation and development or as a result of physiological stresses that engender adaptive responses. We propose that such biochemical signals trigger pathways that ultimately alter DNA binding or the transcriptional activation function of proteins that recognize transcriptional regulatory elements within individual genes or sets of genes. Our experimental goal is to identify the components of such regulatory pathways.

Myoglobin

The myoglobin gene was an attractive choice for our initial studies. Its genomic structure is relatively simple and conserved among mammalian species. It is expressed exclusively in muscle tissues in early experiments (nuclear run-on experiments), we determined that its cell-specific expression is determined by transcriptional activation rather than by post-transcriptional mechanisms involving RNA processing or stability. Furthermore, the activity of the myoglobin gene varies widely among specialized muscle subtypes (Figure 4, top panel). Cardiomyocytes and oxidative skeletal muscle fibers express high concentrations of myoglobin mRNA, resulting in corresponding abundant quantities of myoglobin, which serves as an essential function in supporting high rates of oxygen transport and use in these tissues. However, only very low concentrations of myoglobin mRNA and protein are expressed in glycolytic skeletal fibers.

In addition to its development regulation, expression of the myoglobin gene is subject to physiological control. Electrical stimulation of the motor nerve (a model of tonic work overload) augments expression of myoglobin mRNA in skeletal muscles to levels equal to those in the heart (Figure 4, bottom panel). There also is evidence that gene expression may be regulated by hypoxia, heme availability, and cyclic AMP.

We first searched for cis-acting control elements necessary for transcriptional activation of the human myoglobin gene during differentiation of skeletal myocytes in culture (Figure 5). A region upstream of the transcriptional start site proved to be essential in this process (Figure 6) and was termed a muscle-specific enhancer (MSE). In subsequent experiments, we have identified several protein-binding sites within the MSE. Our studies of the functional importance of these sites, alone and in combination, are still in progress, but our preliminary findings suggest rather complex functional interactions.

Mutation of these protein-binding sites alters MSE function in two ways. Deletion of only a few bases within one protein-binding sequence (motif 1) sharply reduces MSE activity. Multiple copies of a second type of protein-binding sequence (motif 2) are present within the enhancer region but can be mutated or deleted without reducing MSE activity. In fact, inactivation of these latter sites results in somewhat greater transcriptional activity in cultured skeletal myotubes. Although these motif 2 sites appear to provide a brake to transcriptional activation by the MSE during myocyte differentiation in culture, ongoing experiments are addressing the possibility that these sites may serve to modulate transcription in response to physiological stimuli in more mature muscle fibers or cardiomyocytes.

Our efforts to identify proteins that bind to these control elements also are still in progress, but our preliminary findings indicate that at least three types of trans-acting factors bind to DNA within the myo-
Fig. 4. Expression of myoglobin messenger RNA (mRNA) in specialized muscle subtypes. Top panel: RNA blot prepared from striated muscles of two adult rabbits and hybridized to a human myoglobin complementary DNA probe. Size markers are indicated. Myoglobin mRNA is abundant in heart and in oxidative (type I) soleus skeletal muscles but not in glycolytic (type IIb) tibialis anterior (TA) skeletal muscles. Bottom panel: RNA blot prepared from rabbit tibialis anterior skeletal muscles and hybridized to a human myoglobin complementary DNA probe. In control state (C), these muscles express low quantities of myoglobin mRNA. However, when subjected to 21 days of motor nerve stimulation (S), expression of myoglobin mRNA is increased more than 10-fold. Reprinted with permission.68
Muscle-Specific Transcription of Other Genes

Myoglobin is one member of a large class of genes that are expressed selectively in skeletal and cardiac myocytes and that include contractile proteins (e.g., \( \alpha \)-actins), proteins involved in high-energy phosphate metabolism (e.g., M-creatine kinase), cell surface receptors (e.g., acetylcholine receptor subunits), and cytoskeletal proteins (e.g., dystrophin). Other investigators have defined, in varying degrees of detail, cis-acting control elements and trans-acting factors that determine the cell-specific expression of genes encoding many other proteins of this class.71–87

The picture that emerges from these studies is far from simple. Many apparently distinctive cis-acting control elements (at least five) are capable of directing selective transcription of reporter genes in myocytes. Some of these elements are shared among more than one gene, and several genes contain more than one type of element. These distinctive cis-acting control...
elements appear to recognize different cognate protein-binding factors. Efforts to purify and clone these factors are in progress in several laboratories.

Perhaps the most remarkable recent achievement in the field of myocyte growth and differentiation was the discovery by several laboratories\(^{88-92}\) of genes that alter the developmental fate of nonmuscle cells (e.g., fibroblasts) so they begin to express the repertoire of proteins that establish the myocyte phenotype. This group of myogenic determination genes has been the subject of intense scrutiny. The chromosomal locus encoding the prototype member of this group, MyoD1, is heavily methylated in nonmuscle cells, thereby preventing inappropriate expression of this muscle determination factor. When certain lines of fibroblasts are grown in culture in the presence of 5-azacytidine, a hypomethylating agent, some of the cells will begin to express the MyoD1 gene and, under conditions favorable to differentiation (low concentrations of mitogenic growth factors), to display the characteristics of skeletal myocytes. When expression of MyoD1 is forced by introducing the gene into fibroblasts under the control of a strong constitutive promoter, the cells are converted into myocytes at an even greater frequency. The details of how the MyoD1 gene accomplishes this remarkable feat have not yet been fully elucidated, but they appear to include binding of the MyoD1 protein, complexed to other nuclear proteins,\(^{93}\) to transcriptional control regions flanking other genes that are expressed exclusively in

**FIGURE 6.** More detailed localization of transcriptional control elements from human myoglobin gene. Construct studied in experiments shown in Figure 5 was mutated by removing segments of DNA as shown, and transfection experiments were repeated with these mutant constructs. Only certain deletions reduced expression of reporter gene (CAT activity), indicating presence of critical control elements within a defined region. Position of this region (MbURE) relative to transcriptional start site (cap), TATA box, and recognition sites for certain restriction endonucleases (e.g., HindIII, and so on) is shown. Reprinted with permission.\(^{67}\)

**FIGURE 7.** Schematic representation of a functional model for transcriptional regulation of human myoglobin gene during muscle differentiation and in response to changes in contractile work that alter demand for oxidative metabolism. Boxes indicate putative cis-acting transcriptional control elements that are binding sites for nuclear proteins (ovals) that act in a combined manner to form an active transcriptional complex. Functional significance of TATA box and muscle-specific enhancer (MSE) regions have been defined experimentally. Other hypothetical (?) control regions include cyclic AMP (cAMP) response (CRE) and heme-activated (HemeAS) elements. Preliminary evidence suggests presence within the MSE of multiple binding sites that recognize several distinct types of proteins. Proteins with regulatory function during muscle differentiation (e.g., MyoD1, myogenin), and biochemical signaling pathways that potentially may transduce effects of contractile activity to the gene are illustrated.
myocytes. The MyoD1 protein increases transcription of its own gene (positive autoregulation) as well as that of other myogenic determination genes (e.g., myogenin). Muscle-specific transcription of genes that do not bind MyoD1 directly may be induced indirectly by MyoD1-dependent activation of genes encoding other transcription factors. The relations between MyoD1 and other myogenic determination genes and the factors we have identified that bind to transcriptional control elements of the human myoglobin gene are under investigation in our laboratory.

In contrast to skeletal myocytes, MyoD and other known members of the family of myogenic determination genes are not expressed in cardiomyocytes, but proteins that serve an analogous function in these cells will presumably be identified. The ability to alter the developmental fate of mammalian cells by inserting one or a limited number of foreign genes raises the eventual possibility of circumventing a fundamental clinical problem in cardiovascular medicine—the inability of cardiomyocytes to regenerate after damage to a portion of the heart.

Conclusions

In the broadest sense, molecular cardiology can be defined as the application of emerging biotechnologies that permit the isolation, analysis, and manipulation of genes and biologically active gene products for the purposes of promoting greater understanding of the normal biology and pathophysiology of the cardiovascular system and of developing novel therapeutic strategies for the diagnosis, prevention, and treatment of cardiovascular diseases. The goal of this component of the Research Advances Series has been to assist cardiologists and cardiovascular investigators who do not have expertise in molecular biology in developing an understanding of basic principles and vocabulary concerning the control of gene transcription, a topic of considerable importance in efforts to use molecular biology for these purposes.

Advances in molecular biology herald an exciting new era for cardiology in which the developmental fate of cells can be altered to improve the function of the failing heart, inherited genetic defects responsible for devastating cardiovascular sequelae are corrected, and atherosclerosis is halted or reversed by therapeutic manipulation of genes modulating the physiology of the vessel wall. Although this era has not yet arrived, it is approaching at an increasing pace.

References

17. Simon MC, Fisch TM, Bence BC, Nevins JR, Heintz N: Definition of multiple, functionally distinct TATA elements, one of which is a target in the hsp70 promoter for E1A regulation. Cell 1988;52:723–729
23. Pfeifer K, Prezant T, Quarante L: Yeast HAP1 activator binds to two upstream activation sites of different sequence. Cell 1987;49:19–27


34. Williams RS, Johnston SA: Gene transfer into terminally differentiated myocytes. Circulation 1989;80(suppl II):II-602


76. Bouvagnet FR, Strehler EE, White GE, Strehler-Page M, Nadal-Ginard B, Mahdavi V: Multiple positive and negative 5'
regulatory elements control the cell-type-specific expression of the embryonic skeletal myosin heavy-chain gene. Mol Cell Biol 1987;7:4377–4389


86. Piette J, Klarsfeld A, Changeux J: Interaction of nuclear factors with the upstream region of the α-subunit gene of chicken muscle acetylcholine receptor: Variations with muscle differentiation and denervation. EMBO J 1989;8:687–694


89. Wright WE, Sasson DA, Lin YK: Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. Cell 1989;56:607–617

90. Edmondson DG, Olson EN: A gene with homology to the myc similarity region of MyoD1 is expressed during myogenesis and is sufficient to activate the muscle differentiation program. Genes Dev 1989;3:628–640


94. Lassar AB, Buskin JN, Lockshon D, Davis RL, Apone S, Hauschka SD, Weintraub H: MyoD is a sequence-specific DNA binding protein requiring a region of myc homology to bind to the muscle creatine kinase enhancer. Cell 1989;58:823–831


**Key Words** • gene expression • gene transcription • transcription factor • myoglobin
Genetic mechanisms that determine oxidative capacity of striated muscles. Control of gene transcription.
R S Williams

Circulation. 1990;82:319-331
doi: 10.1161/01.CIR.82.2.319

Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 1990 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/82/2/319.citation

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org//subscriptions/