NUCLEOSOMAL STRUCTURE AND FUNCTIONS:
CHARACTERIZATION OF THE HAMSTER CARDIAC MYOSIN HEAVY CHAIN GENES
DNASE I HYPERSENSITIVE SITES

by

Weei-Yuarn Huang

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Clinical Biochemistry
University of Toronto

©Copyright by Weei-Yuarn Huang, 1997
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-27662-7
Nucleosomal Structure and Functions: Characterization of the Hamster Cardiac Myosin Heavy Chain Genes DNase I Hypersensitive Sites
Weei-Yuarn Huang
Department of Clinical Biochemistry
University of Toronto

ABSTRACT

In order to investigate the role of chromatin structure in cardiac gene expression, DNase I hypersensitive sites (DHS) of the hamster cardiac myosin heavy chain genes (MyHC) were studied. Two cardiac-specific DHSs within the 5 kb upstream region of the β-MyHC gene were identified. One of the DHSs was mapped to -2.3 kb (β-2.3 kb) region and the other to the proximal promoter region. Although these two sites were readily detectable using nuclei from neonatal hamster hearts, the proximal promoter site disappeared when adult hamster heart nuclei were used, and the β-2.3 kb site decreased in intensity. Furthermore, the disappearance of the β-promoter DHS in fetal heart chromatin can be induced by injecting thyroid hormone into pregnant hamsters at late gestational stage. Digestion of nuclei from various tissues by micrococcal nuclease revealed that the β-MyHC gene promoter exists in an array of specifically-positioned nucleosomes only in fetal heart chromatin. The β-MyHC gene promoter is DNase I hypersensitive within one of the nucleosomal particles.

As well, two DHSs, one mapping to -1.9 kb (α-1.9 kb) region, the other to the proximal promoter region, were detected within the 4 kb upstream region of the cardiac α-MyHC gene using adult heart
nuclei. It was found that the β-2.3 kb site is associated with multiple conserved muscle regulatory motifs. In the α-1.9 kb site, a conserved GATA motif was identified. The specific interaction between these upstream DHSs and cardiac nuclear proteins was established using gel mobility shift assays and footprinting analysis.

Transient transfection CAT (chloramphenicol acetyl-transferase assays) revealed that (1) the β-2.3 kb site is not a typical enhancer, and (2) the proximal promoter DHS (-294 to +97) of the β-MyHC gene is able to confer tissue-specific expression of the reporter gene; however, (3) co-transfection of GATA-4 with β-MyHC promoter CAT construct into fibroblasts leads to a marked increase in CAT activity. In conclusion, this study demonstrated that developmental regulation of cardiac MyHC gene expression involves the interplay between chromatin structure and cardiac nuclear factors.
ACKNOWLEDGEMENTS

This thesis is dedicated to my parents.

I would like to express my gratitude to the Department of Clinical Biochemistry for offering me the Connaught Scholarship for my first year of graduate studies. It is this generous support that has allowed me to study abroad.

Throughout the course of this study Dr. C. C. Liew has provided me with encouragement and guidance, for which I am indebted. I also wish to thank Drs. H. P. Elsholtz and F. W. Keeley for serving on my supervisory committee, and Dr. Phil Marsden for his critical comments on the thesis.

Finally, I would like to thank N. L. Shih, Eva Cukerman, Ma Hong, Jack Liew, David Hwang, Isolde Prince and other colleagues in Dr. C. C. Liew's laboratory for their advice and assistance, and Margaret for her support during the final stage of the preparation of this thesis.

This work was supported by the Heart and Stroke Foundation of Ontario and the Medical Research Council of Canada.
TABLE OF CONTENTS

Title Page ................................................................. i
Abstract ........................................................................ ii
Acknowledgements ........................................................ iv
Tables of Contents ........................................................ v
List of Figures ............................................................... viii
List of Abbreviations .................................................... x

INTRODUCTION ................................................................. 1
Chromatin Structure and Gene Expression ....................... 2
Nucleosomes, the Unit Particles of Chromatin .................. 3
Some Nucleosomes Are Specifically Positioned on DNA ...... 4
Nuclease Hypersensitive Site ........................................... 6
Nature of the DNase I Hypersensitive Sites ....................... 9
Preset and Remodelling Chromatin Structure ................... 11
Mechanisms of Chromatin Remodelling ........................... 13
DNA Methylation on Chromatin Structure ........................ 22
Chromatin Structure of Developmentally Regulated Genes ... 25
Role of Locus Control Regions ........................................ 27
"Naked" DNA versus Nucleosomal DNA.......................... 28
Transcriptional Regulation of the Cardiac Myosin Heavy Chain Genes Revisited ........................................ 30
Rationale, Hypothesis, and Objectives for the Study ........... 39

MATERIALS AND METHODS .................................................. 41
Isolation of Nuclei and Treatment of DNase I ................. 42
Micrococcal Nuclease Digestion of Nuclei ........................ 42
Preparation of Hamster Cardiomyocytes ........................... 43
Indirect-end Labelling Method ........................................ 44
Preparation of Nuclear Extracts ....................................... 44
Oligomer Construction and Probe Preparation ................ 45
Gel Mobility Shift Assays .............................................. 45
Methylation Interference Analysis ................................... 47
Screening Human Heart cDNA Library ............................. 47
RESULTS

(A) Identification and Characterization of the Cardiac β-MyHC Gene DHS Sites

Cardiac-Specific DHS Sites of the Cardiac β-MyHC Gene

Developmentally-Specific DNase I Hypersensitive Sites of the Cardiac β-MyHC Gene

Thyroid Hormone Causes Disappearance of the β-MyHC Gene Promoter DHS Site

The β-MyHC Gene Promoter DHS Site Is Located on a Specific-Positioned Nucleosome

In Vivo Footprinting Analysis of the β-MyHC Promoter DHS Site

(B) Characterization of the β-2.3 kb Site

Defining the Region of the β-2.3 kb Site

Identification of DNA Binding Activity in the β-2.3 kb Site

(C) Identification and Characterization of the Cardiac α-MyHC Gene DHS Sites

Tissue- and Developmentally-Specific DHS Sites of the α-MyHC Gene

Defining the Region of α-1.9 kb Site

A Conserved GATA Motif Residing within the α-1.9 kb Site

(D) Functional Characterization of β-MyHC Gene
DHS Sites.......................................................93

The Promoter DHS Site of the β-MyHC Gene
Confers Its Tissue-Specific Expression.................93

The β-2.3 kb Site Is Not a Classical Enhancer
in Transient Transfection Assays.......................94

GATA-4 Activates β-MyHC Gene Proximal Promter
in Cardiofibroblasts...................................94

(E) Molecular Cloning of Human GATA-4 cDNAs........103

Isolation and Characterization of
the Human GATA-4 cDNA Clones.......................103

Northern Blot Analysis of the hGATA-4................104

DISCUSSION...................................................107

Chromatin Remodelling of the Cardiac β-MyHC Gene....108

Limitation of DHS Sites Approach........................111

DNase I Sensitive Domain................................112

Cardiac Hypertrophy.....................................114

Chromatin Remodelling and Cardiac Hypertrophy........116

DHS Sites and Familial Hypertrophic Cardiomyopathy...118

The Role of Thyroid Hormone and Its Receptor
in the Formation of DHS Sites.........................119

The Integrity of Cardiomyocyte Nuclei..................122

The β-2.3 kb Site.........................................122

The α-1.9 kb Site.........................................124

GATA-4 and Chromatin Remodelling.......................125

Conclusions................................................128

Future Experiments.....................................129

REFERENCES.................................................134

APPENDIX....................................................179

A1. Chromosome Localization of the hGATA-4 Gene
by in situ Hybridization...............................180

A2. Chloramphenicol Acetyl Transferase Assay Counts...183

A3. Northern Blot Analysis of the Cardiac MyHC Genes..185

VII
Fig. 19 A conserved GATA motif residing within the α-1.9 kb site..................................92

Fig. 20 The promoter DHS site is able to confer tissue-specific expression of the chimeric CAT construct.........................................................95

Fig. 21 Relative CAT activity of the β-2.3 kb site in Thymidine kinase (TK) promoter construct........97

Fig. 22 Transactivation of the β-MyHC gene by GATA-4.........100

Fig. 23 The nucleotide sequence of the human GATA-4 cDNA and deduced amino acid sequence.................105

Fig. 24 Northern blot analysis of hGATA-4....................106

Fig. 25 Chromatin remodelling of the cardiac β-MyHC gene.....................................................133

Fig. 26 FISH mapping of the human GATA-4 gene..................182

Fig. 27 Northern blot analysis of the cardiac MyHC genes....185
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa</td>
<td>amino acid(s)</td>
</tr>
<tr>
<td>ANP</td>
<td>atrial natriuretic peptide</td>
</tr>
<tr>
<td>BHK</td>
<td>baby hamster kidney</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>BNP</td>
<td>brain-type natriuretic peptide</td>
</tr>
<tr>
<td>BrdU</td>
<td>bromodeoxyuridine</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>°C</td>
<td>degree Celsius</td>
</tr>
<tr>
<td>cAMP</td>
<td>cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>CAT</td>
<td>chloramphenicol acetyl transferase</td>
</tr>
<tr>
<td>CBP</td>
<td>CREB-binding protein</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>cpm</td>
<td>counts per min</td>
</tr>
<tr>
<td>CREB</td>
<td>cAMP response element binding protein</td>
</tr>
<tr>
<td>DNS</td>
<td>DNase I hypersensitive site</td>
</tr>
<tr>
<td>DMS</td>
<td>dimethyl sulfate</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EGTA</td>
<td>ethylene glycol-bis(β-aminoethyl ether) tetraacetic acid</td>
</tr>
<tr>
<td>FBS</td>
<td>fetal bovine serum</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>G</td>
<td>gravitational force</td>
</tr>
<tr>
<td>GArC</td>
<td>G, A-rich, C</td>
</tr>
<tr>
<td>GMSA</td>
<td>gel mobility shift assay</td>
</tr>
<tr>
<td>GH</td>
<td>growth hormone</td>
</tr>
<tr>
<td>GR</td>
<td>glucocorticoid receptor</td>
</tr>
<tr>
<td>HAT</td>
<td>histone acetyltransferase</td>
</tr>
<tr>
<td>HEPES</td>
<td>N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid</td>
</tr>
<tr>
<td>hGATA-4</td>
<td>human GATA-4</td>
</tr>
<tr>
<td>hr.</td>
<td>hour</td>
</tr>
<tr>
<td>HSE</td>
<td>heat shock factor response element</td>
</tr>
<tr>
<td>HSF</td>
<td>heat shock transcription factor</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>kDa</td>
<td>kiloDalton</td>
</tr>
<tr>
<td>LM-PCR</td>
<td>ligation-mediated polymerase chain reaction</td>
</tr>
<tr>
<td>LTR</td>
<td>long terminal repeat</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>MEF-2</td>
<td>myocyte specific-enhancer factor 2</td>
</tr>
<tr>
<td>min</td>
<td>minute</td>
</tr>
<tr>
<td>mL</td>
<td>millilitre</td>
</tr>
<tr>
<td>MLC-2</td>
<td>Myosin light chain 2</td>
</tr>
<tr>
<td>mM</td>
<td>millimolar</td>
</tr>
<tr>
<td>MMTV</td>
<td>mouse mammary tumour virus</td>
</tr>
<tr>
<td>MNase</td>
<td>micrococcal nuclease</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger ribonucleic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>MyHC</td>
<td>myosin heavy chain</td>
</tr>
<tr>
<td>N-CoR</td>
<td>nuclear receptor corepressor</td>
</tr>
<tr>
<td>NF1</td>
<td>nuclear factor 1</td>
</tr>
<tr>
<td>ng</td>
<td>nanogram</td>
</tr>
<tr>
<td>nt</td>
<td>nucleotide</td>
</tr>
<tr>
<td>NURF</td>
<td>nucleosome remodeling factor</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethylsulphony fluoroide</td>
</tr>
<tr>
<td>poly(dI-dC)</td>
<td>polydeoxyinosinic-deoxycytidylic acid</td>
</tr>
<tr>
<td>PPET-1</td>
<td>preproendothelin-1</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RSC</td>
<td>remodel structure of chromatin</td>
</tr>
<tr>
<td>RT</td>
<td>reverse transcriptase</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecylsulfate</td>
</tr>
<tr>
<td>SSC</td>
<td>sodium chloride and sodium citrate</td>
</tr>
<tr>
<td>SWI/SNF</td>
<td>switch/sucrose non-fermenting</td>
</tr>
<tr>
<td>TAF</td>
<td>TATA-binding protein associated factor</td>
</tr>
<tr>
<td>TBP</td>
<td>TATA-binding protein</td>
</tr>
<tr>
<td>TH</td>
<td>thyroid hormone</td>
</tr>
<tr>
<td>TK</td>
<td>thymidine kinase</td>
</tr>
<tr>
<td>TPA</td>
<td>12-0-tetradecanoylphorbol 13-acetate</td>
</tr>
<tr>
<td>TR</td>
<td>thyroid hormone receptor</td>
</tr>
<tr>
<td>Tris</td>
<td>tris-hydroxymethylaminomethane</td>
</tr>
<tr>
<td>UAS</td>
<td>upstream activating sequence</td>
</tr>
<tr>
<td>μg</td>
<td>microgram</td>
</tr>
<tr>
<td>μm</td>
<td>micron</td>
</tr>
</tbody>
</table>
INTRODUCTION
Chromatin Structure and Gene Expression

Gene expression in eukaryotes is subject to multiple levels of control. Transcription initiation has been the focus of intense research, and much progress has been made in vitro in understanding this aspect of gene regulation (for review see McKnight, 1995; Struhl, 1996; Orphanides et al., 1996). The first step in transcription initiation is binding to the TATA element by TFIID, a complex containing the TATA-binding protein (TBP) and approximately ten TBP-associated factors (TAFs) (Tjian and Maniatis, 1994). Subsequently, to initiate accurate transcription a set of basic transcription factors (TFIIB, TFIIE, TFIIF, TFIIF, and TFIIIA), and RNA polymerase II assemble stepwisely on a promoter. Experiments using in vitro models that employ "naked" DNA as templates, have shown that gene-specific regulatory factors, which frequently bind to locations upstream of the initiation site, are also required for effective control of transcription. Unlike the naked DNA used in in vitro transcription experiments, eukaryotic DNA is not freely accessible in vivo, but is complexed with histone and non-histone proteins to form a chromatin fiber. It previously was thought that the primary role of chromatin was to act as a passive repressor in transcription. However, recent biochemical and genetic evidence indicates that chromatin is a dynamic structure and may participate actively in gene regulation. Upon gene activation, chromatin structure is changed through the rearrangement of histones and DNA. Eukaryotic gene expression can be understood only in the context of these changes (Svaren and
Nucleosomes, the Unit Particles of Chromatin

A major advance in our understanding of chromatin structure came in 1974 (Kornberg, 1974), with the discovery of the nucleosome, the fundamental packing unit. The nucleosome gives chromatin a beads-on-a-string appearance in electron micrographs taken after treatments that unfold higher-order packing (Richmond et al., 1984; for review see McGhee and Felsenfeld, 1980). The nucleosome "bead" can be removed from the long DNA "string" by digestion for a short period with micrococcal nuclease (MNase). After digestion, only the DNA between the nucleosome beads is degraded. The rest is protected from digestion and remains as double-stranded DNA (dsDNA) fragments (146 bps) bound to a specific complex of 8 nucleosomal histones consisting of 2 copies of each of the H2A, H2B, H3 and H4. This histone octamer forms a protein core around which the dsDNA helix is wound twice (See Fig. 1 for illustration). In undigested chromatin the DNA extends as a continuous thread from nucleosome to nucleosome. Each nucleosome bead is separated from the next by a region of linker DNA, which can vary in length between 0 and 80 nucleotide pairs.

The adjacent nucleosomes are further packed together with the help of H1. In the presence of this molecule, two full turns of the DNA (166 nucleotide pairs) are protected from micrococcal nuclease digestion (Thomas et al., 1979; for review see Pederson, 1986). H1 has an evolutionarily conserved globular central region linked to
extended amino-terminal and carboxyl-terminal "arms" (Cole, 1983; Coles et al., 1987). The globular core of H1 binds to each nucleosome near the site where the DNA helix enters and leaves the histone octamer. Its arms are thought to extend to contact other sites on the histone cores of adjacent nucleosomes, so that the nucleosomes are pulled together into a regular repeating array (Allan et al., 1980) (Fig. 1A). However, the precise sites of interaction of its extended amino-terminal and carboxyl-terminal arms with the nucleosome remains unknown.

In living cells, chromatin rarely adopts the extended "beads-on-a-string" form. Instead, the nucleosomes are packed upon one another to generate higher-order structure (e.g. 10-nm and 30-nm chromatin fiber) (Belmont et al., 1987). How the nucleosomes are arranged in these structures remain a mystery. This introduction will be focused on the studies of chromatin structure at the level of the nucleosome.

Some nucleosomes are specifically positioned on DNA

The positions of nucleosomes in cells can be analyzed by treating cell nuclei with micrococcal nuclease. This enzyme preferentially cuts the linker DNA, and the sites that are protected from digestion can then be analyzed. Regulatory elements of several genes were found to be organized with specifically positioned nucleosomes (Simpson, 1986; Travers, 1987). Such a unique structure has been indicated as an intrinsic part of the regulatory mechanism for eukaryotic genes. For example, in the
Fig. 1 **Nucleosome, the basic unit of chromatin** (A) "beads-on-a-string" form. When H1 is present, two full turns of the DNA (166 bps) are protected from MNase digestion. (B) The octameric histone core consists of a H$_3$H$_4$ tetramer and two H$_2$A.H$_2$B dimers. In a symmetrical model for the nucleosome, the H$_3$.H$_4$ tetramer provides a kernel for the shape.
Drosophila melanogaster hsp26 gene promoter, heat shock factor (HSF) and GAGA factor, bound approximately 200 bp apart, are brought into proximity with each other by the wrapping of the intervening DNA around a specifically positioned nucleosome (Thomas and Elgin, 1988; Lu et al., 1993). The mouse mammary tumor virus (MMTV) promoter (Richard-Foy and Hager, 1987; Truss et al., 1995) and the yeast PHO 5 gene promoter (Almer and Horz, 1986; Schmid et al., 1992) are also packaged into a continuous array of specifically positioned nucleosomes. Gene activation is often accompanied by perturbations of the arrays of nucleosomes that make regions of promoters and enhancers more sensitive to nuclease digestion (i.e., the formation of nuclease hypersensitive site) (Wallrath et al., 1994).

Nuclease hypersensitive site

In chromatin, nuclease hypersensitive sites are believed to represent "open windows" that allow enhanced access of crucial cis-acting DNA sequences to trans-acting factors. These accessible regions are defined by their pronounced sensitivity to nuclease cleavage (such as DNase I, restriction enzymes, micrococcal nuclease, or endogenous nuclease) or chemical modification, and are typically two orders of magnitude more sensitive than are other regions in bulk chromatin (for review see Gross and Garrard, 1988). The method most commonly used to map nuclease hypersensitive sites is DNase I digestion of nuclei followed by indirect end-labelling of the resulting purified dsDNA (Wu, 1980; Fig. 2). However, the
Fig. 2 Indirect-end labelling method DNA in bulk chromatin was first partially digested by DNase I. After purification, genomic DNA was further digested by EcoRI and fractionated by an agarose gel. If the DHS sites are present in the restriction fragment (EcoRI-EcoRI), subbands will be observed in the subsequent Southern blot analysis using the probe, whose position is adjacent to an EcoRI site.
Fig. 3 Ligation-Mediated PCR (LM-PCR) This protocol consists of 4 steps: (A) primer extension (using the first gene-specific primer), (B) ligation with the common linker, (C) PCR amplification (using the second gene-specific primer and top oligomer of the common linker as forward and reverse primers, respectively), (D) nt sequence analysis (using the end-labelled third gene-specific primer). See Materials and Methods for details.
resolution of this method is crude, and these regions are now being elucidated by alternative mapping procedures such as ligation-mediated PCR (LM-PCR, Fig. 3) which yields single-nucleotide resolution (Mueller and Wold, 1989).

A wide variety of functional sequences are associated with hypersensitive sites, including promoters, upstream activation sequences (UAS), enhancers of active or inducible genes (Liu et al., 1988), silencers of transcription (Baniahmad et al., 1990), origins of replication, recombination elements, and structure sites within or around telomeres and centromeres (tabulated see Gross and Garrard, 1988). Additional sites have been mapped for which no function is readily apparent; however, a genetic assessment of such sites has frequently yielded interesting results. For example, two recent studies have demonstrated that some DNase I hypersensitive sites within the human CD2 or β-globin gene locus control region function in the establishment of an open chromatin structure, although these sites lack typical enhancer functions in transient transfection assays (Festenstein et al., 1996; Ellis et al., 1996).

**Nature of the DNase I hypersensitive sites**

Showing that active sites are DNase I-hypersensitive does not tell us much about the nature of these sites. Indeed, the exact mechanisms for hypersensitivity and its role in gene regulation are only partially understood at present. The view held until recently, and based on low resolution nuclease-digestion experiments, was that these sites represent nucleosome-free regions of DNA (Almer
and Horz, 1986; Elgin, 1988; Becker, 1994). For example, using yeast PHO5 gene as a model, Almer and Horz (1986) found that the regular nucleosomal DNA pattern was interrupted in the upstream nuclease hypersensitive sites, and concluded that these nuclease hypersensitive sites represent gaps within ordered nucleosomal arrays. However, this view has recently been challenged. Using high resolution in vivo footprinting analysis, two research groups reported the persistence of nucleosomes within the DHS sites (McPherson et al., 1993; Truss et al., 1995). McPherson et al. (1993) demonstrated that an albumin distal enhancer exists in an array of three precisely positioned nucleosomes only in liver chromatin, where the enhancer is active. This enhancer is hypersensitive exactly between the micrococcal nuclease cleavages that define one of the nucleosomal particle. As well, Truss et al. (1995) demonstrated that hormone induction in MMTV promoter causes a DNase I hypersensitive site limited to a short region in the centre of nucleosome B; this nucleosome is neither removed nor shifted upon hormone induction. The localized nature of the alteration, together with the micrococcal nuclease digestion results which revealed an intact nucleosomal structure, suggest that hormone induction leads to a nucleosome rearrangement rather than to the generation of a nucleosome-free region over the MMTV promoter.

That the DHS site is probably free of nucleosomes was demonstrated by several in vitro chromatin assembly experiments (Workman and Kingston, 1992; Felsenfeld, 1996). When GAL4 (a yeast
transcription factor) bound to purified nucleosome cores, a ternary complex was created, containing the GAL4 derivative and all four core histones. However, when these complexes were challenged with non-specific competitor DNA, they dissociated either into the original nucleosome cores or into complexes of GAL4 bound to naked DNA. About 40% of the products comprised complexes of GAL4 bound to DNA, demonstrating nucleosome displacement as a consequence of GAL4 binding and histone displacement onto competitor DNA (Felsenfeld, 1996). Because of limitations of methodology, such events have not been able to be demonstrated in vivo.

Architectural Variations of Inducible Eukaryotic Promoters: Preset and Remodelling Chromatin Structures

Inducible genes have two types of chromatin structure: preset and remodelling (Wallrath et al., 1994). Preset genes are those in which the binding sites for trans-acting factors are accessible prior to activation (i.e., in a DNase I hypersensitive configuration). In response to the activation signal, trans-acting factors bind to cis-acting regulatory elements and trigger transcription with no major alterations in the chromatin structure of the promoter region. By contrast, remodelling genes are those in which some of the required cis-acting regulatory elements are packaged into nucleosomes (i.e., in a non-DNase I hypersensitive configuration). In order for the trans-acting factors to gain access to cis-acting elements, the nucleosomes must be perturbed in response to an activation signal, a chromatin remodelling
process which forms DNase I hypersensitive sites must occur.

The Drosophila melanogaster hsp26 gene is an excellent example of an inducible preset gene. Prior to heat shock, there are two prominent DHS sites 5' of the transcription start site, which map to the locations of the two functional heat shock response elements (HSEs) (Thomas and Elgin, 1988; Cartwright and Elgin, 1986). Two stretches of alternating C and T residues immediately adjacent to the HSEs bind GAGA factor in vitro (Lu et al., 1993), and are footprinted in vivo (Gilmour et al., 1989). Upon heat shock induction, heat shock factor (HSF) binds to the HSEs, and RNA polymerase II is allowed to continue transcription past the pause site (Rougvie and Lis, 1990; Thomas and Elgin, 1988). No major changes in the chromatin structure upstream of the transcription start site are observed upon activation.

The best studies of remodelling promoters have been carried out in mammalian and yeast systems. In mammalian cells, the MMTV LTR (long terminal repeat) promoter region has an array of six positioned nucleosomes (Richard-Foy and Hager, 1987). One of the nucleosomes (designated nucleosome B) is positioned over the sites to which the glucocorticoid receptor (GR) and the NF1 positive regulators bind (Pina et al., 1990). Immediately upon dexamethasone induction, a DHS site localized at the center of this nucleosome becomes apparent, suggesting a perturbation of the nucleosome. NF1 can now bind to its site, which had been before hormone induction, in-accessible in vivo (Pina et al., 1990; Archer et al., 1991). In yeast, the PHO5 promoter is packaged into six precisely positioned
nucleosomes (Fascher et al., 1990; Schmid et al., 1992). One of the nucleosomes (designated -2 nucleosome) covers one of the two binding sites for the PH04 positive regulator and a binding site for the PH02 positive regulator. In the presence of phosphate, the PH080 gene product is complexed with the PH04 protein, rendering it inactive. Upon phosphate deprivation, PH080 and PH04 dissociate, and PH04 binds to its sites on the promoter. Four of the nucleosomes are perturbed, as demonstrated by DNase I hypersensitivity over the entire promoter region and by the loss of the nucleosomal ladder.

The alteration of chromosomal structure observed in both systems is not due to the consequences of ongoing transcription. In the MMTV promoter, the formation of the DHS site in nucleosome B is unaffected by concentrations of α-amanitin that completely inhibit transcription (Truss et al., 1995). As well, deletion of the PH05 TATA box, which abolishes transcription, does not affect the remodelling process (Fascher et al., 1993). Thus, perturbation of nucleosomal structure appears to a prerequisite for, or an initial step in, gene activation.

Mechanisms of chromatin remodelling

(1) Transcription factors bind to nucleosomes

An important event in the remodelling process appears to be the initial binding of a positive regulator to its site over a nucleosome (Wallrath et al., 1994). For example, in the MMTV promoter, the initial step in nucleosome structure alteration is
the binding of the glucocorticoid receptor (GR) to an intranucleosomal site (Archer et al., 1991). However, not all transcription factors can bind to their sites on the face of a nucleosome. While the GR can bind to a nucleosomal site with almost the same affinity as it binds to free DNA, the NF1 and Oct 1 transcription factors cannot bind to the positioned nucleosome over MMTV promoter before hormone induction (Archer et al., 1991; Archer et al., 1992). It is unclear whether this difference in binding ability is due to the locations of the binding sites of nucleosome or to distinct properties of the proteins. However, the GAL4 and HSF appear to be mechanistically distinct (Taylor et al., 1991), that is, the HSF can bind only to free DNA, whereas, the GAL4 can bind to randomly positioned or rotationally phased nucleosomes in vitro.

(2) Modification of histones

Histones may be modified in two ways: by alterations in histone composition (such as depletion of linker histone H1 and histone H2A/H2B dimers), or by chemical modifications (e.g., histone hyperacetylation). It has been reported that hormone induction results in partial depletion of linker histone H1 from MMTV promoter (Bresnick et al., 1992). Since histone H1 contacts the DNA not only at the entry and exit sites of the nucleosome but also at the pseudo dyad axis (the center of the nucleosome), removal of histone H1 should increase DNA accessibility in the central region of the nucleosome. Dissociation of histone H2A/H2B
dimers together with the formation of a H3/H4 tetramer has been reported, and such a tetramer structure is able to facilitate transcription factor binding to nucleosomal DNA in vitro (Camerini-Otero et al., 1976; Jackson, 1990). For example, *Xenopus borealis* TFIIIA binds with unimpaired affinity to a positioned nucleosome organized in a histone tetramer structure, whereas a positioned nucleosome with a histone octomer structure inhibits binding of TFIIIA (Hayes and Wolffe, 1992). As well, hyperacetylation of the histones can alter nucleosomal conformation (Bauer et al., 1994), as a consequence to destabilization of inter-nucleosomal contacts (Garcia-Ramirez et al., 1995), and to increases in the accessibility of nucleosomal DNA to transcription factors (Lee et al., 1993). The elimination of histone acetylation is correlated with transcriptional silencing (Braunstein et al., 1993).

Recently, several coactivators have been identified as histone acetyltransferases (Brownell et al., 1996; Ogryzko et al., 1996; Yang et al., 1996; Bannister and Kouzarides, 1996; Mizzinen et al., 1996), providing direct evidence that acetylation plays a role in gene expression (Wolffe and Pruss, 1996). Initially, Brownell et al. (1996), purified a subunit of a *Tetrahymena* nuclear histone acetyltransferase (HAT). Molecular characterization of this enzyme reveals a surprising identity to GCN5p, a yeast protein involved in transcription with unclear functions when it was first identified. Subsequent experiments demonstrated that GCN5p itself is a histone acetyltransferase. In earlier studies, GCN5p had been shown to be a regulatory molecule that facilitates the yeast acidic
activators such as GCN4 and GAL4-VP16. GCN5p functions as a complex with two other proteins, the coactivator ADA2p and ADA3p (Georgakopoulos and Thireos, 1992; Georgakopoulos et al., 1995). ADA2p interacts directly with the acidic activation domain of VP16 and with TATA-binding protein (TBP). Therefore, with the discovery that GCN5p is a HAT, it was hypothesized that when a particular set of transcription factors bind to their DNA-binding sites, this regulatory complex ADA2p-ADA3p-GCNSp could be recruited to the vicinity of promoters, resulting in histone hyperacetylation (i.e., nucleosome perturbation) (Wolffe and Pruss, 1996).

The identification of Gcn5p as a histone acetyltransferase led to the discovery that three other mammalian coactivators, p300/CBP, P/CAF and TAF\(_{II}\) 250, are also histone acetyltransferases (Ogryzko et al., 1996; Yang et al., 1996; Mizzen et al., 1996). Although Gcn5p has a limited range of activation, p300/CBP has been shown to be associated with many transcription factors (for review see Janknecht and Hunter, 1996), indicating that acetylation is important for turning on a wide range of genes. Moreover, the discovery that TAF\(_{II}\) 250 subunit of TFIID has histone acetyltransferase activity, suggests that TATA-binding protein is able to overcome the repressive effect of chromatin by directly recruiting one of its subunits to modify chromatin structure.

Like the histone acetyltransferases, the histone deacetylases have also been found to be important for gene regulation. For example, the recent purification and molecular characterization of a human deacetylase has revealed a remarkable identity with the S.
cerevisiae transcriptional regulator, Rpd3p, a global transcriptional regulator required for target genes to maximize transcriptional efficiency in yeast (Taunton et al., 1996). Thus, by a mechanism involving the targeting histone acetyltransferases or deacetylases to a particular gene, chromatin structure could be reversibly modulated to activate or silence transcription.

(3) Recruitment of SWI-SNF complex (or other nucleosome remodelling factors)

Several chromatin remodelling factors have been identified, including SWI-SNF complex (SWI for switch, SNF for sucrose non-fermenting), GAGA factor, NURF (nucleosome remodelling factor), and RSC (remodel the structure of chromatin). The yeast SWI-SNF complex is a conserved multisubunit complex consisting of 11 proteins (for review see Winston and Carlson, 1992; Peterson and Tamkun, 1995). Initially, five subunits of the S. cerevisiae SWI-SNF complex, encoded by the SWI (ADR6), SWI2 (SNF2), SWI3, SNF5 and SNF6 genes, were identified in genetic studies as positive regulators of two genes, HO and SUC2 (Stem et al., 1984; Neigeborn and Carlson, 1984). Subsequently, these regulators were found to be required for the transcription of many other diversely regulated genes (Laurent et al., 1990; Peterson and Herskowitz, 1992). Furthermore, many transcription factors, including the GAL4, require one or more SWI-SNF products to enhance transcription in S. cerevisiae. The activity of GAL4 is reduced at least tenfold in the absence of either SWI1, SNF5 or SNF6 (Laurent et al, 1990). However, SWI-SNF
products are not required for basal transcription in vivo (Peterson and Herskowitz, 1992) or in vitro (Cairns et al., 1994). So far, only ten SWI-SNF-dependent genes have been identified.

A connection between the function of this SWI-SNF complex and chromatin was established through genetic studies. Mutations in several chromatin components alleviate the defects in growth and in transcription that are due to SWI or SNF mutations (for review see Winston and Carlson, 1992). For example, deletion of one of the two gene clusters that encode histones H2A and H2B (Hirschhorn et al., 1992), or missense mutations in genes encoding histones H3 or H4, restores activator function in the absence of SWI or SNF products (Kruger et al., 1995). More recently, biochemical experiments employing a purified complex have provided further support for the idea that SWI-SNF complex is involved in regulating nucleosomal organization. This complex can facilitate the in vitro binding of GAL4 to sites on the surface of a nucleosome in a strictly ATP-dependent manner (Cote et al., 1994). Even in the absence of GAL4, the SWI-SNF complex has a distinct effect on the nucleosome in that the normal 10bp periodicity of DNase I cleavage of nucleosome DNA is markedly perturbed. Although histones remain associated with the DNA, the SWI-SNF complex interferes with histone-DNA contacts, and this is presumably the mechanism by which it assists GAL4 binding to the nucleosome. One of the subunits of the SWI-SNF complex, SWI2-SNF2, has been characterized as a DNA-stimulated ATPase and may be the motor of the nucleosome remodelling machinery (Laurent et al., 1993). Mutations that
inactivate the ATPase activity of SWI2-SNF2 also interfere with its activity in vivo (Cote et al., 1994).

Wilson et al. (1996) recently reported the intriguing findings that SWI-SNF proteins are also integral components of the *S. cerevisiae*’s RNA polymerase II holoenzyme, which contains RNA polymerase II, a subset of general transcription factors and SRB (suppressor of RNA polymerase B) regulatory proteins (Koleske and Young, 1994; Barberis et al., 1995). Since the RNA polymerase II holoenzyme is required for many yeast genes in vivo (Thompson and Young, 1995), it is proposed that the activator binds to its DNA-binding sites and recruits the holoenzyme, and the SWI-SNF components of the holoenzyme enhance the stability of the activator-DNA interaction by destabilizing nucleosomes (Wilson et al., 1996). Although whether the SWI-SNF complex is associated with RNA polymerase II holoenzyme remains an issue of controversy (Cairns et al., 1996), this model is attractive in that it provides a simple solution to the question of how SWI-SNF proteins are brought to promoters (Peterson, 1996).

A second yeast chromatin-remodelling complex (RSC for remodel the structure of chromatin), was identified recently (Cairns et al., 1996b). Like SWI-SNF, RSC exhibits a DNA-dependent ATPase activity and a capacity to perturb nucleosome structure. At least three RSC subunits are related to SWI-SNF polypeptides: Sth1p, Rsc6p, and Rsc8p, and are similar to Swi2/Snf2p, Swp73p, and Swi3p, respectively. RSC is, however, at least 10-fold more abundant than SWI-SNF complex and is essential for mitotic growth, indicating
that RSC plays a wider role in gene expression. No association of RSC with RNA polymerase II holoenzyme was detected.

Soon after the SWI-SNF proteins were identified in S. cerevisiae, evidence began to accumulate that related proteins might assist DNA-binding regulatory proteins in activating transcription in higher eukaryotes. For example, the glucocorticoid receptor (GR) requires SWI-SNF proteins to activate transcription in S. cerevisiae, and antibodies against SWI3 interfere with the ability of GR to activate transcription in Drosophila nuclear extracts (Yoshinaga et al., 1992). Initially, most potential SWI-SNF homologues that have been identified in higher eukaryotes were related to the SWI2 subunit (because of its conserved nature among species), including the brahma (brm), an activator of Drosophila homeotic genes (Tamkun et al., 1992), and the human BRG1 and hbrm (Khavari et al., 1993; Muchardt and Yaniv, 1993; Chiba et al., 1994). The purification and cloning of mammalian SWI/SNF complexes was recently achieved by Wang et al. (1996a and b). The subunits that they cloned, BAF190 (BRG1 associated factors, 190 kD), BAF170, BAF155, BAF60, and BAF47 are similar to yeast SNF2/SWI2, SWI3, SWI3, SWP73, and SNF5, respectively. While the observed similarity between mammalian and yeast complexes is significant, the complexes in mammals are more diverse than the SWI/SNF complex in yeast. The subunit BAF60 is encoded by a novel gene family. Its yeast homology, SWP73, is required for GR activation in yeast (Cairns et al., 1996a). Three members of this family with different tissue-specific expression patterns were cloned, BAF60a, BAF60b, and
BAF60c. While BAF60a is expressed at similar level in various tissues examined, the BAFb and BAFc are predominantly expressed in pancreas and muscle (including heart and skeletal muscle), respectively. The relatively selective expression of BAF60 homologues in different tissues, suggests the existence of tissue-specific chromatin remodelling activity in mammals.

Using a highly refined in vitro nucleosome assembly system from Drosophila, Wu and co-workers characterized the GAGA factor (Tsukiyama et al., 1994) and its ATP-utilizing cofactor, NURF (Tsukiyama and Wu, 1995). GAGA factor is an abundant protein that binds to GA/CT-rich sites upstream of many Drosophila genes (Wallrath et al., 1994). It was found that introducing recombinant GAGA factor during or after nucleosome assembly resulted in a disruption of chromatin structure surrounding the hsp70 promoter (Tsukiyama et al., 1994). This disruption was characterized by DNase I hypersensitivity at the TATA and heat-shock transcription factor binding sites, similar to the nuclease hypersensitivity observed at the endogenous hsp70 gene in nuclei from heat-shocked cells. This hypersensitivity appeared to be caused by a localized rearrangement of nucleosomes surrounding the hsp 70 promoter. In addition, micrococcal nuclease was able to invade and cleave the DNA within nucleosome core particles that contained GAGA binding sites, suggesting that nucleosome structure might be altered. Furthermore, as with SWI/SNF disruption, the GAGA factor-induced effects required ATP hydrolysis, and its ATP hydrolysis ability is derived from its cofactor, NURF (Tsukiyama and Wu, 1995). Purified
NURF is composed of four major polypeptides and sediments as a protein complex of 500 kDa. It uses the energy of ATP hydrolysis to stimulate the binding of GAGA factor to a nucleosomal binding site (Tsukiyama and Wu, 1995). Like SWI-SNF, this reaction also involves the ATP-dependent disruption of histone-DNA contacts by NURF. However, the ATPase activity of NURF is distinct from SWI-SNF; NURF ATPase activity is stimulated only by nucleosomal DNA, not by naked DNA or by free histone (Tsukiyama et al., 1995). By contrast, the ATPase activity of the SWI-SNF complex is stimulated equally by free DNA, nucleosomal DNA, and structured DNA (Cote et al, 1994). These cofactor requirements suggest that the targets, and thus the mechanisms, of nucleosome disruption may be distinct from SWI-SNF and NURF complexes.

**Effect of DNA Methylation on Chromatin Structure**

DNA methylation occurring also plays an important role in the regulation of gene expression during development (for review see, Edan and Cedar, 1994). Most of the methyl groups are found in the C residue of CG "doublets," (5'-'CpG-3") and, in fact, the majority of the CG sequences of eukaryotic genomes are methylated in adult stage. At first, most of the mouse embryo DNA is unmethylated at the blastula stage. Following implantation of the embryo, the bulk of the genome is modified by a wave of de novo methylation activity, leaving unmethylated all of the CpG islands associated with housekeeping genes (Kafri et al., 1992). At later stages of embryogenesis, demethylation occurs in the tissue-specific genes.
in the cell types in which the genes are expressed. As a result of these events, the final adult modification pattern is characterized by full methylation of those genes that are inactive and under-methylation of those genes that are active (Monk et al., 1987; Naveh-Many and Cedar, 1981).

DNA methylation at CpG residues suppresses transcription by affecting DNA-protein interaction (Edan and Cedar, 1994). DNA methylation produces a reversible protrusion from the major groove of DNA and is, therefore, capable of altering local protein recognition signals. Such interference in binding has been demonstrated biochemically for a number of known proteins, such as AP-2 (Comb and Goodman, 1990), c-Myc/Myn (Prendergast et al., 1991), E2F (Kovesdi et al., 1987) and NF-kappaB (Bednarik et al., 1991), all of which recognize sequences that contain CpG residues. However, this is not the case for all transcription factors; several ubiquitous transcription factors, including Sp1 (Holler et al., 1988) and CTF (Ben Hattar et al., 1989) are not affected by DNA methylation.

In addition to its direct effect on factor binding, DNA methylation may also alter overall chromatin structure, thereby indirectly influencing gene accessibility (Keshet et al., 1986). For example, when a skeletal α-actin promoter/CAT chimeric construct was stably introduced into mouse L cells (a fibroblast cell line) in an unmethylated form, the whole integrated construct became DNase I sensitive and the α-actin promoter became DNase I hypersensitive. By contrast, when the same cell line containing the
same construct but with methylation at every CpG residues was generated for analysis, the entire integrated sequences became DNase I-insensitive. No DNase I hypersensitive sites could be detected in the methylated construct. These result suggest that DNA methylation may affect the formation of active chromatin (Keshet et al., 1986).

The inhibitory action of DNA methylation may also require the presence of chromatin structure (Buschhausen et al., 1987). In Buschhausen’s study (1987) naked DNA containing the TK gene promoter was injected into cells, both methylated and unmethylated templates were, initially, transcribed equally. However, once the cells entered the S phase and the injected DNA became packaged into a chromatin structure, the methylated construct was rendered inactive. When the DNA injection is carried out using preformed chromatin templates, the inhibitory effect of methylation was immediate. These results suggest that DNA methylation may serve as a signal that directs the local formation of inactive chromatin structure.

In the above studies, active or inactive chromatin was defined by the presence or absence of DHS sites. The exact nature of chromatin structure in the methylated or un-methylated form remains unknown. After in vitro reconstitution, there was no significant difference between methylated or unmethylated chromatin, as tested by electron microscopy and micrococcal nuclease digestion (Buschhausen et al., 1987). In vivo experiments, however, indicated that nucleosomes are distributed preferentially over methylated DNA
in the genome (Ball et al., 1983; Solage and Cedar, 1978).

DNA methylation may also play a role in genomic imprinting, both as a means for transmitting allele-specific signals from gamete to embryo, and as a cis-acting mechanism for maintaining the imprinted transcription pattern in each cell (for review see Razin and Cedar, 1994). The above-described mechanisms whereby DNA methylation affects transcription factor (or repressor) binding and alters chromatin structure may also be applied to explain part of this epigenomic phenomenon.

Chromatin Structure of the Developmentally Regulated Genes

When gene switching occurs as an intrinsic part of a developmental process, (as in the different forms of globin synthesized in embryonic, fetal, and adult red blood cells), a switch in the pattern of DHS sites associated with different promoters is observed. The presence of these particular 5′-DHS sites correlates with gene activity (McGhee et al., 1981; Stadler et al., 1980b). The appearance of the promoter DHS site in the chicken adult 8-globin gene is a relatively late step in the progression of chromatin structure changes associated with gene activation (McGhee et al., 1981). However, the formation of this DHS site occurs prior to the initiation of gene transcription, as shown by experiments in which chick cells are transformed with a temperature-sensitive avian erythroblastosis virus (Nakajima et al., 1988). In that study, one arrested cell line (due to the temperature-activated erythroblastosis virus) was recovered that
exhibited DHS sites at the β-globin promoter but did not show transcription. Transcription occurred only when the developmental block induced by the virus was released by a temperature shift, allowing the remaining critical step(s) in the activation process to proceed. Many other studies also support the conclusion that formation of developmentally specific DHS site precedes or accompanies gene expression (Elgin, 1988).

Recently, Wood and his collaborators (Stanworth et al., 1995) have performed experiments with somatic cell hybrids (consisting of mouse erythroleukemia (MEL) cells and erythroblasts from transgenic mice carrying part or all of the human β-globin locus), to dissect the molecular mechanisms underlying the developmental regulation of γ- to β-globin switching. The switching of the human globin genes within the transgenic mouse embryo is similar to the process that occurs in the human. When mouse embryonic cells expressing human fetal γ-globin gene were fused with MEL cells (which express only the adult mouse β-globin genes), the human fetal γ-globin mRNA production was maintained, suggesting that the environment of the MEL cells was not sufficient to overcome the information on the chromosome carrying the human genes. Furthermore, if the mouse embryonic erythroid cell was irradiated before fusion, (which resulted in hybrids containing only small fragments of donor chromosome), the pattern of gene expression did not differ from that of unirradiated hybrids. This observation suggests that continued expression of trans-acting factors from the donor erythroblasts is not necessary for continued expression of
the human γ-globin gene in MEL cells. By contrast, when the DNA constructs containing the human globin gene cluster were transfected into MEL cells, an unregulated expression of the globin genes was observed, characterized by a mixed expression of both adult and fetal isoforms (Morley et al., 1991). Thus, when the same constructs were introduced as naked DNA into MEL cells, they behaved quite differently from when they were introduced as intact chromatin, which had been through a normal developmental history. These results indicate that chromatin structure established during normal development is required for developmental stage-specific expression of β-globin gene cluster. (Stanworth et al., 1995; Felsenfeld, 1996)

Role of Locus Control Regions

Locus control regions (LCR) are cis-acting elements that play a crucial role in the formation and maintenance of open chromatin. They are able to confer the transgene's expression in a copy number-dependent and a position-independent manner (Felsenfeld, 1992). In the chromatin of the human β-globin gene cluster, the LCR is marked by a set of four nuclease hypersensitive sites (HS1 to HS4) located 10-20 kilobases upstream of the genes (Grosveld et al., 1987; Forrester et al., 1987; Tuan et al., 1985). If this LCR is coupled to a human β-globin gene (with its promoter and other local regulatory elements) and introduced into transgenic mice, the amount of human β-globin mRNA synthesized is proportional to the number of gene copies that have been integrated (i.e., copy number-
dependent effect). This suggests that the LCR allows gene expression independent of the chromatin environment at the site of integration (i.e., position-independent effect). Without the LCR, integration of the 8-globin gene leads to low levels of expression unrelated to copy number (Grosveld et al., 1987). Although LCR has an enhancer-like effect, it is in most cases distinguishable from canonical enhancer activity. Of the four human 8-globin LCR domains (the four nuclease hypersensitive sites), three do not function as enhancer in transient transfection assays, but only when stably integrated into the genome (Tuan et al., 1989; Talbot et al., 1990 and 1991, Philipsen et al., 1990), suggesting that the activity of the LCR is associated with chromatin structure.

Globin LCR domains are replete with DNA-binding sequences for well-known trans-acting factors (Talbot et al., 1990). Thus, it has been suggested that a primary role for the LCR is to enter into cooperative binding interactions that keep (or induce) the promoter that is free of histones (Felsenfeld, 1992). This hypothesis was further supported by the observation that a globin LCR element alone can be nuclease hypersensitive in erythroid cells of transgenic mice, but that a globin promoter is not nuclease hypersensitive except in the presence of an LCR element (Felsenfeld, 1992).

"Naked" DNA versus Nucleosomal DNA

Using the mouse mammary tumor virus (MMTV) promoter as a model, Archer et al. (1992) demonstrated that there are differences
in gene regulation between a transiently transfected promoter and a promoter organized into a nucleosomal structure. Both classes of MMTV promoter (i.e., transiently transfected promoter and stably transfected promoter) are strongly hormone inducible; however, the degree of induction from the stable promoter was three- to fourfold greater than that from the transient construct, suggesting that the transient promoter may be derepressed (i.e. a low degree of activity was observed before hormone induction) (Croston et al., 1991). In addition, the gene’s characteristic nucleosomal repeat pattern can be detected only in the stable promoter (but not in the transient promoter). The nucleosome structure was found to exclude the transcription factor NF1/CTF from the promoter before glucocorticoid treatment. Upon treatment with glucocorticoids, this phased nucleosomal structure was disrupted, allowing the NF1 to bind to the promoter. By contrast, when the promoter was transiently introduced into cells, NF1/CTF was bound constitutively. Therefore, Archer et al. (1992) proposed that glucocorticoid receptor (GR) activation is a bimodal process. The first step involves a transition in chromatin structure as a consequence of receptor binding, which renders the NF1 site available for occupancy. In a second step, GR interacts with basal transcription factors (because hormone-mediated activation is still observed in transient assays). In this model, NF1 serves as an amplifier, cooperative with GR and leading to synergistic transactivation of the MMTV promoter.

Moreover, GR, NF1, and Oct-1 have been shown bound to naked
DNA in a complex compatible with steric hindrance (Bruggemeier et al., 1990): when NF-1 occupies its binding site, GR and Oct-1 cannot bind to their receptive binding sites on MMTV promoter. However, wrapping of the DNA in the nucleosome imposes a radial orientation of the bound proteins that could reduce steric hindrance, and allow all transcription factors to bind to nucleosomal DNA simultaneously and to activate gene expression synergistically (Truss et al., 1995). Indeed, similar results were obtained in studying HIV-1 transcription in vitro with chromatin template (Pazin et al., 1996). In that purified chromatin assembly system, it was found that synergistic transactivation of NF-κB with Sp1 and HIV-1 enhancer-binding factors (LEF-1, Ets-1, and TEF-3) were achieved with chromatin, but not with non-chromatin templates.

Transcriptional Regulation of the Cardiac Myosin Heavy Chain Genes Revisited

Myosin heavy chains (MyHC), the major contractile proteins in mammalian hearts, are expressed in two isoforms: α-MyHC and β-MyHC (for review see Morkin, 1993; Mably and Liew, 1996). These two isoforms have different ATPase activities (with the α-MyHC harbouring higher ATPase activity and the β-MyHC lower activity) that correlate with the velocity and force of contraction of the cardiac myofibers.

Three cardiac myosin isoforms, V1, V2, and V3, have been identified according to their relative electrophoretic mobilities on a pyrophosphate gel (Hoh et al., 1978). V1 and V3 contain
homodimers of αα- and ββ-MyHCs, respectively, while V2 contains an αβ-MyHC heterodimer. Differences in MyHC composition result in biochemical variation in these myosin isoforms; the relative ATPase activity of V3 : V2 : V1 is 2 : 4 : 7 (Pope et al., 1980).

The mammalian cardiac α- and β-MyHC genes are tandemly aligned in the genome (i.e., β-α) and separated by a 4 kb intergenic region (Mahdavi et al., 1984; Liew et al., 1990; Epp et al., 1993) (see Fig. 7). The cardiac MyHC genes have long served as genetic models for studying cardiogenesis and for elucidating the molecular mechanisms underlying the transcriptional regulation of cardiac-specific gene expression (Boheler et al., 1991; Chien et al., 1993).

The cardiac MyHC genes are expressed in a tissue- and developmentally-specific manner and are primarily regulated at the transcriptional level (Lompre, 1984; Boheler et al., 1992). In the cardiac ventricle of rodents, such as the mouse, β-MyHC transcripts predominate during embryonic and fetal development. At about the time of birth, a switch in cardiac MyHC gene expression takes place, characterized by the down-regulation of β-MyHC gene expression and a concomitant induction of the α isoform. The adult atrium and slow-twitch skeletal muscles continue to express the α- and β-MyHC genes, respectively. Regulation of these genes is also influenced by pathological processes. For example, cardiac hypertrophy causes the reexpression of the β-MyHC gene in the ventricles of small mammals (Waspe et al., 1990; Izumo et al., 1987).
The developmental expression pattern of the human cardiac MyHC genes is different from that of the rodents (Liew et al., 1990; Epp et al., 1993; Morkin 1993). The β-MyHC gene is the predominant myosin isozyme expressed in both human fetal and adult ventricular myocardium. By contrast, the human α-MyHC gene is expressed in atria. Thyroid hormone strongly suppresses expression of the β-MyHC in the rat (Lompre et al., 1984) and the rabbit (Morkin, 1993), but the extent to which this occurs in the human heart is unknown.

Some cases of familial hypertrophic cardiomyopathy (FHC) have been attributed to mutations in the β-MyHC coding sequence (Tanigawa et al., 1990; Anja et al., 1990). Tanigawa et al. (1990) reported in one kindred that an α/β MyHC hybrid gene is co-inherited with FHC. This hybrid gene is composed of α-MyHC sequences fused to β-MyHC sequences; the transition from α to β sequences occurred in exon 27 of the β-MyHC gene. It was proposed that the hybrid gene was produced by an unequal crossover event occurring between the α- and β-MyHC genes. Several missense mutations were identified in exons of the β-MyHC gene (Anja et al., 1990; Elstein et al., 1992). The mechanisms by which these mutations cause the phenotype of FHC remain unknown.

The Syrian hamster cardiac MyHC genes have similar expression pattern as those of other small rodents (see Appendix, Fig. 27, Wang 1994). So far, in small rodents, the complete cardiac MyHC genes DNA sequence is only available in Syrian hamster (Wang 1994). Among the cloned cardiac MyHC genes of various species, the cloned hamster cardiac β-MyHC gene has the longest 5’ upstream region (12
kb) (Wang et al., 1995). Interestingly, one strain of this animal exhibits an autosomal recessive dilated cardiomyopathy (Roberds et al., 1993). The cardiac MyHC genes are expressed in this pathological condition in a manner similar to the gene expressions found in other hemodynamically overloaded models (Venkatakrishnan et al., 1979). Thus, the hamster cardiac MyHC genes become a suitable genetic model for further investigation of normal and pathological heart development.

(1) The cardiac α-MyHC gene

The α-MyHC mRNA is expressed between 7.5 to 8 days postcoitum in the cardiac tube of embryonic rodent. Its expression diminishes as the ventricular chambers form, 8 to 9 days postcoitum and remains at a low level until just after birth (Lyons et al., 1990; Lompre et al., 1984). The α-MyHC gene is developmentally and hormonally regulated at the transcriptional level by a complex interplay of multiple elements. In addition to the well-established thyroid hormone (T₃) response elements (TRE, Gustafson et al., 1987; Flink and Morkin, 1990), the expression of the α-MyHC gene can be influenced by the DNA-binding sites (located within the proximal promoter) for MEF-2 (Molkentin and Markham, 1993), myogenin factors (Molkentin et al., 1993), M-CAT binding factor (TEF-1) (Molkentin et al., 1994), and GATA-4 (Molkentin et al., 1994).

Transgenic mice with mutations in several DNA-binding motifs have been generated. A murine α-MyHC/CAT gene construct containing
4 kb of the $\alpha$-MyHC upstream region is sufficient to drive both developmental and tissue-specific expression (Subramaniam et al., 1991). While mutations in T$_3$ response elements reduced the expression of the $\alpha$-MyHC/CAT constructs significantly, the mutation in the MEF-2 DNA binding site does not affect the transgene's expression (Subramaniam et al., 1993; Adolph et al., 1993).

Regulation of the $\alpha$-MyHC gene by GATA-4 was demonstrated by directly injecting adult rat myocardium with the $\alpha$-MyHC/CAT constructs containing mutations in GATA sites, resulting in 88% decrease in CAT activity (Molekentin et al., 1994). GATA-4 is a cell-specific transcription factor, belonging to a group of related zinc-finger proteins that recognize the consensus GATA motif (WGATAR, W:A/T, R:A/G) (Arceci et al., 1993). Vertebrate GATA factors contain a conserved DNA-binding domain, which is composed of two related zinc fingers of the general form CXNCX17CNXC. GATA-1 has been shown to regulate the expression of a variety of genes in erythroid cells (Orkin, 1992). GATA-2 regulates the expression of the preproendothelin-1 (PPET-1) gene in endothelial cells (Lee et al., 1991). By gene targeting in mouse embryonic stem cells, GATA-2 has been shown to play a critical role in hematopoiesis (Tsai et al., 1994). GATA-3 is highly expressed in T cells and embryonic brain cells, and has been implicated in the regulation of $\alpha$- and $\delta$-T cell receptor gene expression (Ho et al., 1991; Ko et al., 1991). GATA-4 mRNA is abundantly expressed in adult heart, with some expression also found in gonads, gut epithelium, and yolk sac endoderm (Arceci et al., 1993; Huang et al., 1995). Using \textit{in situ}
hybridization and immuno-histochemistry, Heikinheimo et al. (1994) found that GATA-4 mRNA and protein are expressed in progenitor cells associated with mouse embryo cardiac development. In addition to the α-MyHC gene, GATA-4 has been shown to regulate the expression of the brain-type natriuretic peptide (Grepin et al., 1994) and cardiac troponin C genes (Ip et al., 1994).

Upstream elements outside the proximal promoter region of the α-MyHC gene have also been suggested as playing a role in its expression. Although the α-MyHC/CAT construct containing the 4 kb upstream region responds appropriately to hyperthyroid and hypothyroid stimulation, the α-MyHC/CAT construct whose distal 2.8 kb region has been deleted, only partially suppress transgene expression in response to hypothyroid stimulation (Subramaniam et al., 1993). The GArC motif locating between -908 to -869 of the human α-MyHC gene was initially characterized by Mably et al. (1993), using DNase I footprinting analysis. However, the in vivo significance of the GArC motif remains to be determined.

(2) The cardiac β-MyHC gene

The regulatory elements identified in the cardiac β-MyHC gene promoters of several mammalian species are summarized in Fig. 4. Of particular interest, two M-CAT motifs, one extending from -267 to -288 (distal enhancer core), and the other from -215 to -196 (proximal enhancer core) of the rat β-MyHC gene have been described by Kariya et al. (1993, 1994). The factor interacting specifically with this motif is the rat homologue of TEF-1 (transcriptional
Fig. 4 Multiple regulatory elements detected in the β-MyHC gene promoter are highly conserved between species. M, hamster; R. rat; H, human
enhancer factor 1), a human transcription factor that was initially identified to bind to a viral enhancer (Farrance et al., 1992; Yockey et al., 1996). It has been suggested that this factor is involved in the α-adrenergic pathway leading to up-regulation of the β-MyHC gene in a cardiomyocyte culture model of cardiac hypertrophy (Kariya et al., 1994).

The critical role of TEF-1 in murine cardiogenesis is demonstrated by a mutation in TEF-1 generated by insertion of a retroviral gene trap vector (Chen et al., 1994). Normal myocardial development is affected, as the homozygous embryos die between E11 and E12, showing an abnormally thinning of both the compact layer and the trabeculae of the ventricular wall. However, the β-MyHC gene and some cardiac-specific genes, such as cardiac troponin C, T, and I, which have been shown to be regulated by TEF-1 in vitro (Alexandre, 1994), are still expressed at relatively normal levels in the mutant hearts.

The thyroid response elements (TRE) responsible for inhibiting β-MyHC expression remain undefined by transient transfection analysis (Morkin, 1993). Although upstream thyroid hormone receptor (TR) binding sites have been identified by gel mobility shift assay and DNase I footprinting with TRs, these sites could be deleted or mutated without affecting T3 regulation (Edwards et al., 1992).

The results from transgenic studies suggest that other, as yet undefined, elements may reside further upstream of the β-MyHC gene. Transgenic analysis with the mouse β-MyHC gene promoter revealed that 5 kb of the upstream sequence is required for high-level
expression, and is able to direct copy number dependent and position independent expression (Rindt et al., 1993; Knotts et al., 1995). Expression driven by a shorter construct (600 bp of the upstream region) was two to three orders of magnitude lower when the CAT activity was compared (Rindt et al., 1993). This 600 bp construct showed position effects in transgenic mice and did not demonstrate copy number dependence, although transgene expression remained muscle-specific (Knotts et al., 1995). Furthermore, while mutating any one of the three cis-acting elements within the proximal promoter (i.e., M-CAT, Sp1, and Be3) of the 600 bp construct significantly affected the level of transgene expression, in the context of the 5 kb upstream region, none of these three elements alone is essential for expression of the gene (Knotts et al., 1994). However, simultaneous mutation of the three sites within the 5 kb promoter construct significantly reduce expression. These data suggest that: (1) sequences upstream of -600 can substitute functionally for any one of these regulatory motifs and are important for high levels of expression, and (2) these three regulatory elements do play an important role in the regulation of the β-MyHC gene.

Consistent with the conclusions of transgenic studies, work done by Sadoshima et al. (1992) suggests that the stretch-response element of the rat cardiac β-MyHC gene lies beyond -628 of the 5' flanking region. In this stretch-induced cardiac hypertrophy model, the regulatory element that responds to stretch was mapped beyond -628 of the 5' flanking region, providing more evidence for the
Rationale, Hypothesis, and Objectives for the Study

Gene regulation in eukaryotes is achieved by the combined action of multiple regulatory proteins and transcription factors interacting with each other on complex arrays of cis-acting elements in promoter and enhancer/silencer regions. There is growing evidence that chromatin structure may actively participate in these processes, partly by modulating access to transcription factor binding sites through changing the arrangement (post-translational modification) of histones and DNA. Although many studies, using naked DNA as template, have identified a few of the transcription factors regulating cardiac-specific gene expression, very few investigator have specifically examined the role of chromatin in cardiac-specific gene expression.

For the past decade, our laboratory has focused on characterizing the cardiac MyHC genes (Liew et al., 1990; Mably et al., 1993; Epp et al., 1993; Wang et al. 1994 and 1995). It is hypothesized that transcriptional regulation of the cardiac MyHC genes can be understood only in the context of changes in chromatin structure. The complete sequencing of the hamster cardiac MyHC genes (Wang et al., 1994 and 1995), allows us to use the DNase I hypersensitive sites as an approach to address the following questions:

(1) Do cardiac MyHC genes contain regulatory elements associated with specific chromatin structure (i.e., genomic
regions that are DNase I hypersensitive)?

(2) Are there changes in the chromatin structure (at the nucleosome level) of the cardiac MyHC genes during cardiogenesis?

(3) What are the transcription factors interacting with these DNase I hypersensitive sites?

(4) What are the functions of these DNase I hypersensitive sites?
MATERIALS AND METHODS
1. Isolation of nuclei from tissues and treatment of DNase I

Nuclei were isolated from freshly excised liver and heart ventricle of adult (3-month-old), neonatal (1-day-old), 3-day-old, and fetal (late gestation) Syrian hamsters (Jackowski and Liew, 1980). The nuclei were suspended in a buffer containing 15 mM Tris (pH 7.5), 60 mM NaCl, 0.5 mM spermidine and 0.15 mM spermine. The nuclei were then treated with DNase I (Pharmacia) at concentrations ranging from 1 μg to 10 μg/ml at 37°C for 2 min in the presence of 5 mM MgCl₂. The nuclei were then disrupted in 1 M NaCl solution containing proteinase K (200 μg/ml) and 0.25% SDS and incubated overnight at 37°C. Genomic DNA was then extracted three times with phenol and chloroform and precipitated with sodium acetate and ethanol.

2. Micrococcal nuclease digestion of nuclei

Freshly prepared nuclei were resuspended in 95 μl nuclear buffer (60 mM KCl, 15 mM NaCl, 5 mM MgCl₂, 0.1 mM EGTA, 15 mM Tris-HCl pH 7.4, 0.5 mM DTT, 0.1 mM PMSF, 300 mM Sucrose). A serial dilution (4, 2, 1, 1/2, 1/4, 1/8 units/μl) of micrococcal nuclease was made freshly in 5μl nuclear buffer plus 0.4 mM CaCl₂. The nuclei were digested by micrococcal nuclease at 25°C for 3 min, and the reactions were stopped by adding one volume stop solution (20 mM EDTA, 1% SDS). The nuclei were then disrupted in 1 M NaCl plus proteinase K (200 μg/ml), and digested overnight at 37°C. The purification of genomic DNA was performed as described above.
3. Preparation of hamster cardiomyocytes, cardiofibroblasts and BHK cells

Neonatal hamster cardiomyocytes were prepared, with minor modifications, as previously described for neonatal rat cardiomyocytes (Wang et al., 1992). One-day-old neonatal hamster hearts were excised and minced in 1xPBS, and the cells were dispersed by successive additions of 1xPBS containing 0.125% trypsin and stirred at 37°C. The cells were then collected in F10 medium + 20% fetal bovine serum and preplated for 45 min to remove fibroblasts. The enriched cardiomyocytes were removed and replated at a density of 1x10^7/100-mm dish with 0.1 mM bromodeoxyuridine (BrdU) to prevent proliferation of non-myocytes. Thirty-six hours later, the cells were harvested by treatment with 0.05% trypsin-EDTA, washed 3 times with 1xPBS, and resuspended with nuclear buffer containing 60 mM KCl, 15 mM MgCl_2, 0.1 mM EGTA, 15 mM Tris-HCl (pH 7.5), 0.5 mM DTT, 0.1 mM PMSF, 300 mM Sucrose, and 0.5% NP-40. After incubation at 4°C for 10 min, nuclei were collected by centrifugation at 1000g, washed three times with nuclear buffer (without NP-40) and resuspended at a concentration of 1x10^7/ml. The isolation of nuclei from cardiac fibroblasts (harvested from the preplating plates) and baby hamster kidney cells (BHK) was carried out as described for the cardiomyocytes. Aliquots of nuclei were digested with DNase I (from 1 μg/ml to 10 μg/ml) and the digested genomic DNA was purified as described in section one.
4. Indirect-end labelling method (Wu, 1980) (See Fig. 2)

Probes used in this study were amplified by PCR, using the cosmid DNA containing the 5' upstream region of the hamster MyHC genes as template. The amplified PCR fragments were subcloned into the TA vector (from Invitrogen) and subsequently verified by dideoxy DNA sequencing using an automated DNA sequencer (Pharmacia). The positions of probe A (β-MyHC gene) and probe B (α-MyHC gene) adjacent to NcoI and SacI restriction sites, respectively, are indicated in Fig. 7. The genomic DNA isolated from section one or three was then digested by Nco I (for β-MyHC gene) or Sac I (for α-MyHC gene) and fractionated on 1% to 1.4% agarose gels. The gel was then transferred to nitrocellulose and hybridized with the probe, which was prepared by random primer labelling. The nitro-cellulose was then washed in 0.1xSSC and 0.1% SDS at 65°C for 20-45 min and exposed to film with intensifying screen at -70°C for 2-7 days.

5. Preparation of nuclear extracts

Nuclear extracts of neonatal hamster cardiomyocytes, cardioblast, and C2C12 myotubes were prepared as described by Andrews and Faller (1991). Briefly, cells were collected in 1.5 ml cold 1xPBS, then pelleted for 10 seconds and resuspended in 400 μl cold Buffer A (10 mM HEPES-KOH pH 7.9 at 4°C, 1.5 mm MgCl₂, 10 mM KCl, 0.5 mM DTT, 0.2 mM PMSF). The cells were allowed to swell on ice for 10 minutes. Samples are centrifuged for 10 seconds, and the supernatant fraction is discarded. The pellet is resuspended in 20-
100 µl of cold Buffer C (20 mM HEPES-KOH pH 7.9, 25% glycerol, 420 mM NaCl, 1.5 mM MgCl₂, 0.2 mM EDTA, 0.5 mM DTT, 0.2 mM PMSF) and incubated on ice for 20 min. Cellular debris was removed by centrifugation for 2 min at 4°C and the supernatant fraction was collected. All extract preparations were aliquoted and stored at -70°C. Protein concentrations were determined using a modified Bradford assay with bovine serum albumin as the standard (Boehringer Mannheim).

6. Oligomer construction and probe preparation

Oligonucleotides were synthesized on a Pharmacia LKB Gene Assembler Plus, and gel purified for the following experiments. Annealing of oligonucleotide probes for gel mobility shift assay was achieved as described by Mueller and Wold (1989). Briefly, the two oligomers of each probe (see the following section for nucleotide sequence) were brought to a final concentration of 20 pmol/µl in 250 mM Tris, pH 7.7, heated to 95°C for 5 min., transferred to 70°C for 10 min., and then slowly cooled (3 hours) to 4°C. The annealed probes were stored at -20°C. Using the Klenow fragment (Pharmacia), 5' thymine overhangs or restriction sites (for probes derived from cloned DNA fragments) could be preferentially labelled at the 3' recessive end by fill-in reaction. Probes were labelled to a specific activity of 10⁷ to 10⁸ cpm/µg.

7. Gel mobility shift assays (GMSA)
The assay was carried out as described by Chodosh (1989). DNA-protein binding reactions were carried out for 15 min at room temperature, in the presence with 10000 cpm of the labelled probes, 10-15μg nuclear extract, 2μg poly(dI-dC) and various unlabelled competitors (as described in the figure legends). The reaction mixtures were subjected to electrophoresis on 4% polyacrylamide gels with recirculation of the low-ionic-strength electrophoresis buffer between chambers and the apparatus cooled by water. The non-specific competitor (145 bp) used in Fig. 16 is derived form the AccI digested cloned β-2.3 kb site in TA vector. The choice of larger DNA fragment as a probe or a competitor is not unprecedented. Previously, in order to identify the nuclear factor that binds to the transcriptional control elements of the immunoglobulin genes, a DNA fragment (300 bp) has been used as a probe in GMSAs (Singh et al., 1986). In that study, a plasmid (3 kb) was used as a non-specific competitor. The nucleotide sequence of probes and various competitors used in different experiments are listed below for clarification:

For Fig. 16:
(1) CT/ACCC box (C, competitor, derived from β-2.3 kb site)
\[5'\text{-GAAGCGGCCCCTCCCTCCAGTCCCTGT-3'}\]
(2) MEF-2 binding site (M, competitor, derived from the mouse creatine kinase gene)
\[5'\text{-CTCGCTCTAAATAACCCCTGT-3'}\]

For Fig. 19:
(1) Hamster α-MyHC gene GATA binding site (probe)
5'-GTCCTGAAGGACATGATAAGGGACTGAGAACAT-3',

(2) PPET-1 GATA-2 binding site (S2, competitor)
5'-CAACGTGCAGCCGGAGATAAGGCCAGGCCGAA, and

(3) MEF-2 binding site (competitor, derived from β-2.3 kb site)
5'-TTCAGTCCCTGTTTTATTTATAGCACTTGGTC-3

For Fig. 22:
(1) B-MyHC gene GATA motif (probe)
5'-gaatgtaagAGATATttttgcttcgct-3'

(2) B-MyHC gene mutant GATA motif (probe)
5'-gaatgtaagACCGCTttttgcttcgct-3'

(3) Competitors: S2 and C (see above)

8. Methylation interference analysis

Investigation of DNA-protein interactions was achieved using the methylation interference assay. Briefly, end-labelled DNA (i.e., the 85 bp DNA fragment used in gel mobility shift assay as a probe in Fig. 16) was methylated with DMS, as described by Baldwin (1989), and used in a binding reaction that was scaled up fivefold. After electrophoresis and autoradiography of the wet gel, bands were excised and electro-eluted, and the DNA was cleaved by piperidine. The cleavage products were separated on denaturing 8% polyacrylamide gels and exposed overnight at -70°C with intensifying screen.

9. Screening human heart cDNA library

Adult human heart total RNA was obtained by the acid
guanidinium thiocyanate-phenol-chloroform method (Chomczynski and Sacchi, 1987). A 246bp cDNA sequence that encodes the conserved zinc finger DNA binding domain was then generated by RT-PCR (Sambrook et al., 1989), using a pair of oligonucleotide primers derived from murine GATA-4 amino acid sequences (forward primer: 5'-GAGTGTGTSAACTGTGGSGC-3'; reverse primer: 5'-TASAGGCCACAGGCRTTGCA-3'; S = C or G, R = A or G). This cDNA fragment was labeled with $^{32}$PdATP by random hexamer priming and used to screen a λgt11 human heart cDNA library (Clontech). All screening, hybridization, and subcloning techniques were performed using standard protocols (Sambrook et al., 1989).

10. Sequencing human GATA-4 cDNA clones

The inserts of the four positive clones hybridizing to the probe were subcloned into M13mp18 and/or M13mp19 vector. The preparation of single-stranded DNA and dideoxy sequencing reactions using T7 DNA polymerase with fluorescent M13 primers were performed as previously described (Epp et al., 1993). The nucleotide sequence was revealed by an automated A.L.F. DNA sequencer (Pharmacia), and analyzed using the HIBIO DNASIS™ analysis program (Hitachi Software Engineering Co., Ltd.).


Total RNA (20 µg) from human heart, brain and liver were subjected to formaldehyde-agarose gel electrophoresis (Sambrook et al., 1989). After transfer, the membrane was baked at 80°C under
vacuum, prehybridized, hybridized with $^{32}$P-labelled hGATA-4 probe (HGA-1, a 1.1-kb EcoRI fragment from λgt11), and washed following a standard protocol (Sambrook et al., 1989). The hybridization condition was 42°C in the presence of 50% formamide, 10x Denhardt's solution, 0.1% SDS, 0.2 mg/ml denatured sheared sperm DNA, 10 mM EDTA, and 25% 4xRNA (3 M NaCl, 0.6 M Tris-HCl [pH 7.5], 0.18 M NaH2PO4·H2O, 0.24 M Na2HPO4, 0.01 M Na pyrophosphate) for 16 hours. The final wash condition was 50°C with 0.2x SSC and 0.1% SDS for 15 minutes. The membrane was exposed for 24 hours at -70°C with intensifying screen.

12. Neonatal rat cardiomyocytes cultures and transfection experiments

Primary cardiomyocyte cultures were prepared from 1-day neonatal rat (Wistar) ventricular myocardium as described previously (Wang et al., 1992). Cells were dispersed by successive additions of 1 x phosphate-buffered saline (PBS) containing 0.125% trypsin and stirring at 37°C. After preplating to eliminate nonmuscle cells, cardiomyocytes were plated at a density of 1 x 10^7 cells/100-mm dish. Cells were maintained in F10 culture media supplemented with 20% fetal bovine serum. The cells were grown for 36 h, then the CAT construct (15 µg) containing the proximal $\beta$-MyHC gene promoter (-294 to +97, relative to the transcription-initiation site of the $\beta$-MyHC gene) and pRSV $\beta$-galactosidase control plasmid (10 µg) were introduced into cells using calcium phosphate precipitation method (Sambrook et al., 1989). After 6 h,
the cells were washed with 1 x PBS and fresh media was added. After another 48 h, the cells were assayed for chloramphenicol acetyltransferase (CAT) and β-galactosidase activities (Gorman et al., 1982; Rosenthal, 1987). For each CAT assay equal amounts of protein were assayed (100 μg), and the reaction were allowed to proceed for 2 h. The CAT assay results were quantitated using the Bio-Rad phospho-imager, and the mean for three transfections was plotted after adjusting for β-galactosidase activity. The CAT assays were repeated with at least three separate preparation of cells.

The C2C12 cell line was grown in Dulbecco’s modified Eagle’s medium (DMEM) with 10% fetal bovine serum. After confluence was reached, the cells were induced to differentiate by switching the culture media to DMEM with 2% horse serum. The cells at various developmental stages are shown in Fig. 5. After 4 days, the cells were transfected with CAT constructs. After another 48 h, cells were harvested for CAT assays as described above. Cardiofibroblasts were derived from the preplating plates of primary cardiomyocyte cultures, and maintained in F10 media with 10% fetal bovine serum. In co-transfection experiments, the pMT-2 GATA-4 expression vector (10 μg, a gift from D. B. Wilson) or pMT-2 control vector (10 μg was transfected with a β-MyHC promoter CAT construct (15 μg, -294 to +97) or a β-MyHC promoter CAT construct with mutant GATA motif into fibroblasts. The subsequent CAT assays were performed as described for cardiomyocytes.
Fig. 5 Differentiation of C2C12 cell line (A) Myoblasts (40x) (B) one day after serum deprivation (grown in 2% horse serum, 100x) (C) Myotubes (100x, 4 days in 2% horse serum) (D) Myotubes (200x)
Fig. 6 Site-directed mutagenesis The transformed clones in site-directed mutagenesis were further analyzed by dideoxy sequencing method. The GATA motif of the hamster cardiac B-MyHC gene (-271 to -266) is mutated in Clones 1 and 2 (from 5'-agatat-3' to 5'-acccgct-3'). Clone 3 contains the normal sequence.
13. Site-directed mutagenesis

Site-directed mutagenesis was performed by the method originally described by Kunkel (1985). Initially, the proximal promoter region of the α-MyHC gene (-294 to +97, PstI- XbaI) was subcloned into TA vector, a phagemid that can produce single-stranded DNA in the presence of helper phanges. The production of single-stranded DNA, followed by annealing with the mutant oligonucleotide (5'-gaatgtaagACCGCTttttgcttcgct-3'), and double-stranded synthesis were performed as previously described (Sambrook et al., 1989). The mutant constructs were then verified by dideoxy sequencing method, and the results were shown in Fig. 6. Insert from one of the verified mutant clones was subcloned into the pCAT-Basic plasmid (Promega) to generate mα-MyHC promoter/CAT construct.

14. Data analysis

Student’s T-test was used to assess differences between transfections of reporter gene constructs and controls in CAT assays. P-values (<0.01) and standard deviation (S.D.) measurements are indicated in the results and/or figure legends.

15. Analysis of chromatin by LM-PCR (ligation-mediated PCR)

The LM-PCR was performed on DNase I cleaved DNA essentially as described (Muller and Wold, 1989). Sequenase (DNA polymerase version 2.0 from U.S. Biochemical) was used for the initial primer extension reaction (oligomer A1), and the reactions were performed at 37°C. The extended products were ligated with the common linker
(Muller and Wold, 1989). In this step, 1 μl of T4 DNA ligase (1 U/μl; Bethesda Research Laboratories) was used per reaction.

Following ligation of the linker oligonucleotide, the DNA was precipitated with ethanol and resuspended in dH2O. Products from the primer extension and linker ligation were subjected to PCR using the second gene-specific oligonucleotide (oligomer A2) and the linker top oligonucleotide as forward and reverse primers, respectively. In the PCR reaction, template DNA was transferred to a 0.5 ml eppendorf tube, in the presence of 10 μl of 10xTaq buffer (Pharmacia), 0.4 mM dNTP solution, 2 μl of 5 pmol/μl linker top oligo, 2 μl of 5 pmol/μl gene-specific oligo 2, dH2O to 100 μl final volume, and 1 U of Taq DNA polymerase (Pharmacia). Samples were covered with mineral oil and subjected to 22 PCR cycles, consisting of 1 min at 94°C, 2 min at the optimal annealing temperature (60°C) and 3 min at 74°C for extension.

Following PCR amplification, 30 μl of the PCR mixture was removed and transferred to a new 0.5 ml eppendorf tube, along with a mixture of 2 μl of 10x Taq buffer, 2 μl of 5 mM dNTP mix, 2 μl of 1 pmol/μl oligomer (A3, phosphorylated with [γ-32P]ATP), 14 μl dH2O, and 1 U of Taq DNA polymerase (Pharmacia). Reactions were covered with mineral oil and subjected to 4 cycles of PCR amplification, consisting of 1 min at 94°C, 2 min at the 64°C, and 3 min at 74°C. After amplification, the reactions were extracted with an equal volume of chloroform. Finally, the reaction products were ethanol precipitated and resuspended in loading buffer, and separated on 6% polyacrylamide, 7 M urea sequencing gels.
The nt sequences of each gene-specific oligonucleotide used are listed as follows:

A1, 5'-GAAACAGAGCCAGGCTAG-3' (-401 to -384, relative to the transcription initiation site of the β-MyHC gene);

A2, 5'-TAGTGCTGGGGGCCAGGACTCCAA-3' (-386 to -362); and,

A3, 5'-AGGACTCCAAGGGGCACATGCCAG-3' (-371 to -351)

16. Maxam-Gilbert sequencing reactions (Sambrook et al., 1989)

Maxam-Gilbert sequencing reactions were run in parallel with the in vivo footprinting reactions so that the nt sequences of the protected region could be determined. Briefly, 5 μg of plasmid DNA (containing the β-MyHC gene promoter, -401 to +97) was subjected to four standard chemical sequencing reactions (i.e., G, G+A, T+C, and C reactions), and followed by piperidine cleavage. The fragmented DNA were then amplified by LM-PCR as described above, consisting of extension with A1 primer, ligation with the common linker, and amplification by PCR, using A2 primer and the linker top oligomer as forward and reverse primers, respectively. The nucleotide sequences were then revealed by Taq DNA polymerase with an end-labelling A3 primer.

17. Animal model with hyperthyroidism

The procedures described were in accordance with the University of Toronto guidelines for animal experiments. Hyperthyroidism was induced in timed-pregnant Syrian hamsters by administration of L-thyroxine (1 mg/kg/day, intramuscularly)
between 7 and 13 days gestation (The gestation period of hamsters is 14 to 15 days). The animals were sacrificed on day 14 of gestation period. Serum levels of free T₄ and T₃ were measured by a standard clinical radioimmunoassay (performed by Dr. P. Y. Wong at Toronto General Hospital, Dep. of Clinical Biochemistry, U of Toronto). The isolation of fetal heart ventricular nuclei and analysis of DNase I hypersensitive sites were as previously described.
RESULTS
1. **Cardiac-specific DHS sites of the cardiac α-MyHC gene**

The structure of the tandemly-linked β- and α-MyHC genes with the positions of probe A (the β-MyHC gene) and probe B (the α-MyHC gene) is shown in Fig. 7. Using probe A, two strong cardiac-specific DNase I hypersensitive sites were detected within the 5 kb upstream regulatory region of the β-MyHC gene (Fig. 8A, lane 5). These sites were not found in liver chromatin (lanes 6 to 8) or in protein-free genomic DNA digested by DNase I (lanes 1 to 2). The appearance of these two DHS sites is dependent on the dose of DNase I concentration, suggesting that optimal DNase I concentration is required for detecting these sites. Based on the size of DNA fragments detected in the gel, and on the position of probe A, as shown in Fig. 7, these two DHS sites were mapped to -2.3 kb and -0.2 kb, respectively. The -0.2 kb DHS site appeared to be located within the promoter region of the β-MyHC gene. Since only 50 to 60% of cells in the neonatal heart are cardiomyocytes, the DHS pattern of enriched cardiomyocytes was compared with another major cell type of the heart, the cardiofibroblasts. As shown in Fig. 8B, the two strong DHS sites were detected in the cardiomyocytes (lanes 1 and 2) but not in the cardiofibroblasts (lanes 3 to 5) or BHK cells (lanes 6 and 7). Less DNase I concentration was used in this experiment to probe the chromatin structure of the cardiomyocytes. This is probably caused by the nuclei isolated from cultured cells.
are purer than those of intact heart tissues (see Discussion). Thus, by using two different methods of cardiac nuclei preparation and by comparing heart chromatin with at least four different controls, the two DHS sites were shown to be cardiac-specific.

Several weaker subbands were also observed in fibroblasts and BHK cells (Fig. 8B, lanes 4 to 7). While some of them are ubiquitous, others are not observed in heart chromatin. The ubiquitous subbands may represent the genomic sequences that are preferentially digested by DNase I (Liu et al., 1988). The causes of subbands in nonmyocytes are unknown. It has been shown that several weaker subbands detected by DNase I turned out to be the linker regions between individual nucleosomes (Almer and Horz, 1986). Thus, the appearance of these non-specific subbands may be due to the unique nucleosomal organization of the MyHC genes in the non-cardiomyocyte genome.

2. Developmentally-specific DNase I hypersensitive sites of the cardiac β-MyHC gene

The patterns of the DHS sites in the cardiac β-MyHC gene during various developmental stages were also studied. Two of the representative blots are shown in Fig. 9. When nuclei isolated from adult hamster hearts were used, the strong DHS site of the β-MyHC gene promoter completely disappeared, and the -2.3 kb site became weaker (Fig. 9A, lanes 1 to 3). The gradual disappearance of this promoter DHS site was demonstrated by comparing nuclei isolated from cardiac ventricle at various developmental stages, including
Fig. 7 Structure of the hamster cardiac α- and β-MyHC genes. The positions of probes A and B are shown. DHS sites are indicated by arrowheads. The positions of the two NcoI (N) restriction sites in β-MyHC gene are -5269 and +1722, respectively. The two SacI (S) sites in the cardiac α-MyHC gene are located at -3915 and +415, respectively. The inset table summarizes the appearance of the DHS sites identified in this study in various developmental stages of the β- and α-MyHC genes.
Fig. 8 Cardiac-specific DHS sites of the β-MyHC gene. (A) Nuclei from neonatal heart (1-day-old, lanes 3-5) or liver (lanes 6-8) were treated with increasing concentrations of DNase I at 37°C for 2 min as indicated at the top of the gel. Purified genomic DNAs (25 μg) were digested with NcoI, electrophoresed on a 1% agarose gel, and subjected to Southern blot hybridization with probe A. Lanes 1 and 2 are protein-free heart DNA digested with 0.05 and 0.01 μg/ml DNase I, respectively, at 23°C for 30 sec.
Fig. 8 (continued) (B) Isolation of nuclei from cardiomyocytes (My, lanes 1-2), fibroblasts (Fb, lanes 3-5) or BHK cells (lanes 6-7) and treatment with various concentrations of DNase I, as indicated, was conducted according to the procedure described in Materials and Methods. Restriction enzyme digestion of purified genomic DNAs (15 µg) and Southern blot analysis with probe A were as described in (A).
late-gestation fetus (Fig. 9B, lanes 1 to 3), 1-day-old hamster (Neo, lanes 4 to 6) and 3-day-old hamster (3d, lanes 7 and 8). Although the intensity of 8-2.3 kb site does not change significantly, the promoter DHS site could no longer be detected in the nuclei of 3-day-old hamster heart cells. The higher background observed in Fig. 9B, may be caused by the lower specific activity of the probe used in this experiment.

It might be argued that the lower intensity of the promoter DHS site observed in the neonatal heart chromatin was caused by the overdigestion of nuclei by DNase I (Fig. 9B, lane 4). However, we thought it unlikely that the neonatal heart chromatin in lane 4 is overdigested. While the intensity of both the major band and the -2.3 kb subband in the neonatal heart chromatin (lane 4) is comparable to the intensity of the fetal heart chromatin at the same concentration of DNase I (lane 1), a decrease of approximately 50% in the intensity of the promoter DHS site was observed in the neonatal heart chromatin (lane 4). The intensity ratios of the major band, the -2.3 kb subband and the promoter subband in lane 4 (the neonatal nuclei) relative to lane 1 (the fetal nuclei) are 1.04, 0.9, and 0.54, respectively. This observation suggests that there exist intrinsic differences in DNase I sensitivity between the fetal and neonatal heart chromatin in the B-MyHC gene promoter region. The overdigestion pattern should be similar to the patterns of lane 2 and lane 3 (Fig. 9B), in which the intensity of all bands decreases proportionally. Furthermore, we have obtained similar results, using different batches of nuclei. (see Discussion)
3. Thyroid hormone causes disappearance of the \( \beta \)-MyHC gene promoter DHS site

The expression of the \( \beta \)-MyHC gene can be decreased by thyroid hormone (Morkin, 1993; Chizzonite and Zak, 1984). Whether the administration of thyroid hormone has any effect on the chromatin structure of the \( \beta \)-MyHC gene was investigated. As shown in Fig. 10, injecting thyroid hormone into timed pregnant hamsters caused the disappearance of the \( \beta \)-MyHC gene promoter DHS site in the ventricular nuclei isolated from late gestational fetus (lanes 3-5). This experiment has been repeated twice, and the representative blot is shown. In this blot, while the change in intensity of the promoter DHS site is significant, it appears that there is no significant change in intensity of the \( \beta \)-2.3 kb site between the treated and control groups. This result indicates that thyroid hormone may have a major effect on the promoter DHS site.

4. The \( \beta \)-MyHC gene promoter DHS site is located on a specifically-positioned nucleosome defined by micrococcal nuclease digestion

To further investigate the chromatin structure of the cardiac \( \beta \)-MyHC gene, late-gestation fetal heart nuclei were digested by micrococcal nuclease (MNase). Electrophoretic analysis of nuclear MNase-digested samples revealed the expected nucleosome repeat length due to partial chromatin digestion (Fig. 11A, lanes 1 to 9). The purified genomic DNA was further analyzed by the indirect-end labelling method. As shown in Fig. 11 (B), a prominent nucleosome
Fig. 9. Developmentally-specific DHS sites of the β-MyHC gene (A)
Each lane represents genomic DNA (25 μg) purified from adult heart (Ad, lanes 1 to 3) or neonatal heart (Neo, 1-day-old, lanes 4 to 6) nuclei digested with increasing concentrations of DNase I as indicated. Experimental procedure from nuclei isolation to Southern blot analysis with probe A were as described in Fig. 8A. Lane 7, liver nuclei digested by 5 μg/ml DNase I.
Fig. 9 (continued) (B) The DHS sites of the β-MyHC gene in late gestation fetal heart ventricle (FH, lanes 1 to 3), neonatal heart ventricle (Neo, 1-day-old, lanes 4 to 6), and 3-day-old heart ventricle (3d, lanes 7 and 8) were compared.
Fig. 10 Thyroid hormone alters the DHS sites of the cardiac β-MyHC gene. DHS pattern of the β-MyHC gene in late gestational fetus treated with thyroid hormone (T+, lanes 3-5) were compared with untreated group (T-, lanes 1-2). Each lane was loaded with 20 μg genomic DNA purified from ventricular nuclei treated with DNase I as indicated. Subsequent Southern blot analysis with probe A was performed as above described. Serum T3 and T4 levels of timed pregnant hamsters are shown in the inset table. T test indicated that both free T4 and T3 levels are significantly different between these two groups (p<0.01).
repeat was observed in liver (Fig. 11B lanes 1 to 4) digested by MNase. According to the position of probe A (Fig. 7) used in this study, this nucleosome repeat was mapped to the transcribed region of the B-MyHC gene.

However, in the late-gestation fetal heart (during which stage, this gene is actively transcribed), the characteristic nucleosome repeat became less prominent in the transcribed region (Fig. 11B lanes 5 to 7). Instead, three specific-positioned nucleosomes (defined by MNase cleavage) around the proximal promoter region were detected only in the late-gestation fetal heart nuclei. It was found that the B-2.3 kb site in fetal heart nuclei is also sensitive to MNase digestion (Fig. 11B lanes 5-7).

The above described experiment was repeated in Fig. 11C, using adult ventricular nuclei and different batches of nuclei preparation including liver (lanes 4-6) and fetal heart ventricle (lanes 7-9). The adult heart chromatin and liver chromatin showed similar nucleosomal organization in the transcribed region of the cardiac B-MyHC gene (Fig. 11C lanes 1-6). As well, around the B-MyHC promoter, an array of specific-positioned nucleosomes was observed in the fetal heart chromatin (lanes 7-9).

Interestingly and reproducibly, inclusion of a DNase I-treated late-gestation fetal heart nuclear sample on the Southern blot showed that the B-MyHC gene promoter is DNase I hypersensitive exactly between the micrococcal cleavages that define the N2 particle (Fig. 11C lanes 7 to 10, also see Fig. 11D). Thus, the N2 particle, if it is nucleosomal (as indicated in this study), may
be perturbed. The position of each nucleosome particle relative to the transcription-initiation site is summarized in Fig. 11D. Since the resolution of MNase cleavage followed by indirect end-labelling analysis is relatively low, the exact position of each nucleosome in this promoter region remains unclear. However, TATA box (-37 to -32) is estimated to be located within the N2 particle.

5. In vivo footprinting analysis of the B-MyHC promoter DHS site

Protein-DNA interaction within the hamster B-MyHC promoter DHS site was investigated by in vivo footprinting analysis. It was expected that the B-MyHC promoter DHS site could be revealed to the nucleotide level of resolution by using the protocol of ligation-mediated polymerase chain reaction (LM-PCR) developed by Mueller and Wold (1989, see Fig. 3 for illustration of this protocol). The result of analyzing the promoter region between -343 to -154 by LM-PCR is shown in Fig. 12. The position of each primer relative to the transcription initiation site of the B-MyHC gene is shown in Materials and Methods. When the cardiomyocyte's genomic DNA from DNase I-cleavage nuclei was amplified, the Be1 element (containing the putative thyroid hormone response elements) and the Be3 element were clearly protected. Several unsuccessful attempts were made to obtain a clearer footprinting of the other regulatory elements within this region. (Our lack of success might have been because the contacts between those nuclear factors and the DNA were disrupted during the preparation of nuclei.)
Fig. 11 Tissue- and developmentally-specific nucleosome positioning over the β-MyHC promoter  

(A, B, and C) Nuclei from liver, adult heart ventricle (AD) and late-gestation fetal ventricle (FH) were treated with increasing concentration of MNase where indicated. (A) Electrophoretic analysis of representative MNase-digested samples revealed the expected nucleosome repeat length due to partial chromatin digestion (lanes 1 to 9). For comparison, nuclear DNase I samples from fetal heart (lanes 10 and 11) were run in parallel.
Fig. 11 (continued) (B and C) Indirect end-label analysis of chromatin samples that were digested with NcoI (20 μg), electrophoresed in a 1.4% agarose gel, transferred to nitrocellulose, and hybridized to the probe A shown in (D). Lane M (B, lane 8; C, lane 11) contains genomic DNA digested with NcoI + PstI (2013 bp, -294) or NcoI + XbaI (1622 bp, +97). Lane D (C, lane 10) contains late-gestational fetal genomic DNA purified from nuclei treated with DNase I (3 μg/ml). The promoter DHS site is indicated by a ""]" at the side of lane 10 (C). The positions of MNase cleavage in various samples are indicated by dots at the side of each lane. A non-specific band at the position higher than the -2.3 kb site is observed in (C) (lanes 6 to 10). The lambda phange digested by Bst EII was used as a DNA marker in (B) and (C).
Fig. 11 (continued) (D) Summary of MNase and DNase I cleavage sites over the β-MyHC promoter in late-gestation fetal heart chromatin. The position of nuclease cleavage are shown by arrows; inferred positions of nucleosomal particles are shown by the closed circles and given the names N1-N3. Because the N2 particle is DNase I hypersensitive and is likely to be a perturbed structure, its circle is shown as a dotted line.
Fig. 12 *In vivo* footprinting analysis of the cardiac β-MyHC gene promoter. Genomic DNA (1 μg) isolated from neonatal heart nuclei (My, lanes 5 and 6) treated with various concentrations of DNase I were amplified using the LM-PCR protocol (see Fig. 3, p. 8 for protocol, and *Materials and Methods*, p. 53, for the position and nt sequence of each primer). As comparison, genomic DNA (Gm, 1 μg) directly digested by DNase I, and subjected to the same LM-PCR protocol were run in parallel (lanes 3 and 4). Lanes 1 and 2 are the Maxam-Gilbert sequencing ladders, using the TA vector containing the β-MyHC gene promoter region as a template (see *Materials and Methods*, p. 55). The DNA-binding motifs identified by *in vitro* footprinting analysis (Be1, Be2, and Be3, see Fig. 4, p. 36), by transient transfection assays (GATA-4, this study), or by transgenic analysis (Sp1/GC rich region) are marked for clarification.
(II) Characterisation of the 8-2.3 kb site

6. Defining the region of the 8-2.3 kb site

To further localize the 8-2.3 kb site, several hamster genomic DNA markers were run in parallel with this site (Fig. 13). Based on the DNA markers, the size of the 8-2.3 kb site was estimated at between 4177 bp and 4032 bp (i.e., -2460 bp to -2315 bp relative to the transcription-initiation site of the 8-MyHC gene). Since the size range of most DHS sites is about 200 bp (Elgin, 1988), the 8-2.3 kb site was defined roughly to a range of 250 bps (-2508 bp to -2258 bp) to include most of the potential regulatory elements residing within this region. It is interesting that comparison of this 250 bp DNA sequence in human and hamster demonstrated 85% identity between the two species, while the sequence similarity outside this region was significantly lower (Fig. 14).

7. Identification of DNA binding activity in the 8-2.3 kb site

As shown in Fig. 14 and 15, multiple regulatory motifs which have been shown to be important for muscle-specific gene expression were identified within the 250 bp region of the 8-2.3 kb site. These include three E-boxes (Blackwell and Weintraub, 1990), one AP-2 site (Buskin and Hauschka, 1989), one CArG motif (Miwa and Kedes, 1989; Treisman, 1992), one CT/ACCC box (Feo et al., 1995), and one MEF-2 site (Gossett et al., 1989) (Fig. 15A). Except for the AP-2 site and one of the E-boxes (-2513 to -2517), other sites are conserved across species. Although this AP-2 site is not
conserved, its nucleotide sequence in human remains quite similar to the consensus binding sites (SSSNKGGG, S: G/C, K: G/T, N: A/C/G/T) for AP-2 factor (Buskin and Hauschka, 1989).

Of particular interest, a 85 bp subregion of this -2.3 kb site contained a MEF-2 site juxtaposed with an E-box motif and a CT/ACCC box. The proximity of a MEF-2 site and an E-box motif within the enhancers or promoters of numerous muscle-specific genes is well documented (Jayne et al., 1988; Nakatsuji et al., 1992; Kaushal et al., 1994). Also, in the enhancers of a few muscle-specific genes, the DNA spacing between the MEF-2 site and CT/ACCC box is a multiple of a DNA helix turn (10.5 bp) (Feo et al., 1995). In our case, a 20 to 22 bp spacing was found between the two core motifs. The enhancer of the mouse MCK gene (-1256 to -1050) (Nakatsuji et al., 1992) contains regulatory elements similar to those identified within the B-2.3 kb site (see Fig. 15B).

It has been suggested that DNA-binding sites on a fragment of 25 to 100 bp can be used as a probe in gel mobility shift assays (GMSA) (Chodosh, 1989). Therefore, GMSAs using this 85 bp fragment as a probe were performed to determine whether cardiac nuclear proteins could interact with binding sites within this region. As shown in Fig. 16A, two specific complexes were identified (lane 2). The complexes could be competed away by cold probe (lane 4) but not by a large excess of an unrelated DNA competitor (lanes 5, 6, a 145 bp AccI-AccI subfragment of the B-2.3 kb site, see Materials and Methods p. 46). Furthermore, when two oligomers (one derived from the CT/ACCC box of the B-2.3 kb site and the other from the MEF-2
Fig. 13 Defining the region of the 8-2.3 kb site  Probe A (Fig. 7) was used in this study. Lane 1, genomic DNA (25 μg) from adult heart nuclei treated with 3 μg/ml DNase I was digested with Nco I and run in parallel with DNA markers (lanes 2 to 5). The -2.3 kb site is indicated by an arrowhead. Lanes 2 to 5 represent 25 μg genomic DNA digested by SphI, PstI, SacI + NcoI, and PstI + NcoI, respectively. The size of DNA fragments from lanes 2 to 5 are 4177 bps, 4032 bps, 3783 bps and 2015 bps, respectively.
Fig. 14 The 3.2-3.3 kb site is highly conserved among species. M: hamster, H: human
binding site of the mouse MCK gene) were added together as a competitor, these two complexes were completely competed away (Fig. 16B, lane 6). However, neither of these two oligomers alone was sufficient to eliminate binding to the 85 bp DNA fragment (Fig. 16B, lanes 2 to 5).

A plausible interpretation of the above observation would be based on the assumption that both factors (CT/ACCC box and MEF-2) cannot simultaneously bind to the 85 bp subfragment (because of the proximity between the CT/ACCC box and MEF-2 binding sites, see Fig. 14). When the CT/ACCC box motif and MEF-2 binding motif of the 8–2.3 kb site was used alone as a probe in GMSA, the intensity of the retarded complexes is similar to that of lane 2 and lane 4 in Fig. 16B, respectively (data not shown), indicating that there is abundant CT/ACCC factor, but not very much MEF-2 present in our cardiac nuclear extracts. Thus, abundant CT/ACCC factor could occupy a large amount of the probe, resulting in the complex A and B observed in Fig. 16, and little competition by the MEF-2 binding site in the lane 5 of Fig. 16B (however, note that the background in lane 5 is still much lower than that of lane 2). Not until in a large excess of CT/ACCC box binding motif was present as a competitor (Fig. 16B, lane 4), was the binding activity of the MEF-2 observed in Fig. 16B (lane 4).

Alternatively, the CT/ACCC factor can bind to the 85 bp probe much more efficiently than can the MEF-2. In the absence of CT/ACCC box competitors, much of the probe was occupied by CT/ACCC factor, and complexes A and B were formed (Fig. 16B, lane 2). In the
presence of a large excess of CT/ACCC box oligonucleotides, most complexes A and B were eliminated (Fig. 16B, lanes 3-4), and the MEF-2 was able to bind to the probe (Fig. 16B lane 4). Thus, the complete elimination of the complexes in Fig 16B (lane 6) is probably due to the competition between both factors for their binding sites on the probe. Such a possibility can be tested by using large amounts of recombinant MEF-2 in a GMSA, to see whether the formation of the complexes will be affected. As well, the nature of the complex formed in lane 4 (Fig. 16B) could be identified by using the antibody against MEF-2 in a GMSA. A supershifted band should be observed in the presence of a large excess of CT/ACCC box competitors.

The relatively close contact between the specific nucleotides and proteins in complexes A and B were determined by methylation interference analysis using the 85 bp fragment as a probe. In this assay, DMS methylates guanine residues at the N-7 position that protrudes into the major groove, and also at the N-3 position of adenines that protrudes into the minor groove. A DNA probe that is methylated at a position which interferes with binding will not be retarded in this assay. Therefore, the specific DNA-protein complex (detected in a gel mobility shift assay) is depleted for DNA that contains methyl groups on purines important for binding. After gel purification, DNA is cleaved with piperidine. Finally these fragments are electrophoresed on polyacrylamide sequencing gels and autoradiographed. Guanines and adenines that are interfered with binding are revealed by their absence in the retarded complex
relative to a lane containing piperidine-cleaved free probe. Thus, comparing lane 3 (Fig. 16C, complex A) and lane 1 (Free DNA) reveals a protected region corresponding to the position of the CT/ACCC box within the 85 bp fragment. The overall lower intensity found in lane 4 (complex B) is because of loading of a lesser amount of DNA. Because of weaker DMS reaction of A residues within the MEF-2 binding site (lanes 1 and 2, Baldwin, 1989), we were unable to demonstrate any footprinting within this region. The protection of AT-rich sequences may be revealed by treatment of the probe with DEPC (Diethylpyrocarbonate) (Molkentin and Markham, 1994).

Weak G residues remained in the protected region of CT/ACCC box binding site (lanes 3 and 4). This observation could be explained by the previous assumption that CT/ACCC box and MEF-2 factor may not be able to bind to the probe simultaneously. Because a minor portion of the retarded complexes are attributed to the MEF-2 binding activity alone, some methylated G residues of the CT/ACCC box within such a portion of the retarded complexes will not interfere with the binding of the CT/ACCC box proteins, resulting in partial protection.
Fig. 15 Nucleotide sequence comparison of the 8-2.3 kb site (A) with the enhancer of the mouse MCK gene (B). Both genes are co-expressed in cardiac and skeletal muscles. The cardiac 8-MyHC gene, however, is only expressed in slow-twitch skeletal muscle (see Introduction, p. 31).
(A) Cp:  
<table>
<thead>
<tr>
<th></th>
<th>S</th>
<th></th>
<th>NS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5</td>
<td>10</td>
<td>200</td>
<td>5</td>
</tr>
<tr>
<td>0</td>
<td>5</td>
<td>200</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(B)  
<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>M</th>
<th>C+M</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>5</td>
<td>200</td>
<td>500</td>
</tr>
<tr>
<td>0</td>
<td>5</td>
<td>200</td>
<td>500</td>
</tr>
</tbody>
</table>

Fig. 16 DNA binding activity in the 8-2.3 kb site. (A) From lanes 2 to 6, 10 μg nuclear extracts from neonatal hamster heart were incubated with an end-labelled AccI-PstI DNA fragment (85 bp, a subregion of the 8-2.3 kb site, see Fig. 15A) at room temperature for 15 min in the presence of nonradioactive specific (cold probe, S) or non-specific (NS, 145 bp AccI-AccI subfragment of the 8-2.3 kb site) competitors (Cp) added at the molar ratios shown. Retarded migration of complexes A and B that represent DNA-binding activity are indicated. (B) Panel B displays competition analysis using CT/ACCC box (C, lanes 2 to 4), or MEF-2 oligomers (M, lane 5) as competitors. In lane 6 both CT/ACCC box and MEF-2 competitors were added together (MEF-2 binding site, 500 molar excess; CT/ACCC box, 500 molar excess).
Fig. 16 (continued) (C) Protein-DNA interactions in complexes A (lane 3) and B (lane 4), were revealed by methylation interference analysis (see: Materials and Methods). Lanes 1 and 2 are free DNA (F). The protected residues (5'-GGAGGGAGG-3', antisense strand) in complexes A and B are shown on the right. The vertical bar indicated the MEF-2 site.
Identification and characterization of the cardiac α-MyHC gene DHS sites

8. Tissue- and developmentally-specific DHS sites of the α-MyHC gene

The 4 kb upstream region of the cardiac α-MyHC gene was also examined. Using probe B, and hybridizing with genomic DNA from adult and neonatal heart ventricles, a strong cardiac-specific DNase I hypersensitive site was detected within the 4 kb upstream regulatory region of the α-MyHC gene (Fig. 17A, lanes 4-10). This site was not found in the genomic DNA extracted from kidney nuclei (Fig. 17A lanes 1 to 3) or liver nuclei (Fig. 17A lanes 11 to 13). Based on the size of DNA fragments detected in the gel, and on the position of probe B, as shown in Fig. 7 (see p. 60), this DHS site was mapped to -1.9 kb (α-1.9 kb site). This site was also detected in the late-gestational fetal ventricle (Fig. 17B, lanes 1 to 3). A very weak subband mapped to the proximal promoter region was barely detectable in adult heart ventricles (Fig. 17A, lanes 5 to 7; Fig. 17B, lanes 6 and 7). The diffuse subband (around 700 bp) detected around the promoter region is probably caused by the long gel electrophoresis and relatively low concentration of agarose (1.3%) used in this study.

9. Defining the region of α-1.9 kb site

Whether the α-1.9 kb site is specific to cardiomyocytes was investigated. As shown in Fig 18A, the α-1.9 kb site is positive
Fig 17 The DHS sites of the cardiac α-MyHC gene  Genomic DNA (25 μg in each lane) purified from nuclei of various tissues treated with increasing concentrations of DNase I, as indicated, were digested with SacI, fractionated on 1.3% agarose gels and subjected to Southern blot analysis with probe B (see Fig. 7, p.60). (A) Ad, adult heart ventricle; Neo (1-day-old), neonatal heart ventricle. Lambda DNA digested by BstEII was used as a DNA marker.
Fig. 17 (continued) (B) The experiment in (A) was repeated using the fetal heart ventricle nuclei (FH, lanes 1 to 3), and different batches of adult heart nuclei (Ad, lanes 4-7). Lambda DNA digested by BstEII was used as a DNA marker.
Fig. 18 Defining the region of the α-1.9 kb site (A) Genomic DNA (25 μg) from cardiofibroblast (Fb, lanes 1-2), and neonatal cardiomyocytes (Myo, lanes 3-4) nuclei, treated with DNase I as indicated, were digested by SacI and run in parallel with DNA markers (lane 5). These markers were generated by digesting genomic DNA with (1) SacI (4330 bps), (2) ApaI (4037 bps), (3) SmaI (2602 bps), and (4) ApaI + SmaI (2382 bps).
Hamster cardiac α-MyHC gene (α-1.9 kb site)

-1972
acagagtatggcttttgctacttggacttgaccccaggctgacccaa
tgttctcag\text{TTCTTATCTGTCC}ttcaggaccttgaaccaggca
GATA motif
gtgacatattaggccacaggctaatccctgtgacttgacacaaggtga
ccttcagggacctagctgcagacaggtggcttgcatctcttgaga
acaatcattttggcatagtcacctgcagatggaatacaaggttgac
tcaggtcccttcaagagaact

-1722

\begin{align*}
\text{(A)} & 5'-' \text{ggtcctgaAGGACATGATAAGGGActgagaac-3'} \\
\text{(B)} & 5'-' \text{agctcctgGGGACATGATAAGGGAGctgagaac-3'} \\
\text{(C)} & 5'-' \text{caccataaGGGACATGATAAGGGGactgca-3'}
\end{align*}

(A) Hamster GATA motif; (B) Human GATA motif

(C) GATA-1 binding site in the Human β-globin gene 3' enhancer

---

Fig. 18 (continued) (B) The nt sequence of the α-1.9 kb site is listed. For comparison, the nt sequences of the human GATA-binding site within the corresponding region of the hamster α-1.9 kb site, and GATA-1-binding site in the human β-globin gene 3' enhancer are provided.
only in the cardiomyocytes (lanes 3-4), and could not be observed in the nuclei from cardiofibroblasts, indicating that the α-1.9 kb site is myocyte-specific.

To further localize the α-1.9 kb site, several hamster genomic DNA markers were run in parallel with this DHS site. According to the DNA markers (Fig. 18A, lane 5), the upper limit of the α-1.9 kb site was estimated to be about 2382 bp. Since the size range of most DHS sites is about 200 bp (Elgin, 1988), the α-1.9 kb site was roughly defined to a range of 250 bps (-1972 bp to -1722 bp) to include most of the potential regulatory elements residing within this region.

10. A conserved GATA motif residing within the α-1.9 kb site

As shown in Fig. 18B, an evolutionarily conserved GATA motif was identified within the α-1.9 kb site. The core motif and flanking sequences of this GATA-binding site are almost identical to that of a well-established GATA-binding site located within the human β-globin gene 3' enhancer (Wall et al., 1988; Fig. 18B). The specific GATA-binding activity in the cardiomyocyte nuclei extracts was revealed by GMSA shown in Fig. 19. This DNA-binding activity can be competed away by the cold probe (lane 3) and the GATA-2 binding site (lane 4) derived from PPET-1 gene (Lee et al., 1991), but it cannot be competed away by the CT/ACCC box (lane 5, C/G rich). When the MEF-2 oligomer was added in large molar ratios, a slight decrease in GATA-binding activity was observed (lane 6). This decrease was most likely because the MEF-2 oligomer used in
this study contained a short stretch of DNA sequence (ttatag) similar to the consensus GATA-binding site (A/TGATAA/G). A band-shift (lane 9) which migrated faster than the GATA-binding activity in lane 2 was observed in the nuclear extracts of the C2C12 myotubes. Since the band-shift can easily be competed away by non-specific competitors (lane 12 and 13), the DNA-binding activity observed might be due to non-specific protein-DNA interaction or to a related GATA factor binding to this site.
Fig. 19 A conserved GATA motif residing within the α-1.9 kb site. Nuclear extracts from neonatal hamster cardiomyocytes (lanes 2 to 7), cardiofibroblasts (lane 8) or C2C12 myotubes (lanes 9-13) were incubated with an end-labelled GATA-binding site (nt sequences see Materials and Methods, p. 46) in the presence of various competitors, as indicated (S1: cold probe, S2: GATA-2 DNA-binding site, C:CT/ACCC box, M: MEF-2). The complexes detected in cardiomyocyte extracts (lane 2) and C2C12 myotube extracts (lane 9) are indicated by an arrow. In order to show the free probe (F), the DNA-binding reaction carried out in lane 2 (cardiomyocytes) and lane 9 (C2C12 myotubes) were loaded into lane 14 and 15, respectively, half an hour later than the other lanes. The complex in lanes 14-15 corresponds to the complex in lane 2 is also indicated by an arrow at right side. Less amount of nuclear extracts was loaded in lane 15.
(IV) Functional characterization of β-MyHC gene DHS sites

11. The promoter DHS site of the β-MyHC gene confers its tissue-specific expression

To define the region of the β-MyHC promoter DHS site, two DNA markers (generated by digestion of genomic DNA with restriction enzymes as indicated in Fig. 20A) were run in parallel with this site. Based on the DNA markers, the range of this site is roughly estimated at between -294 to +97 relative to the transcriptional-initiation site.

A transient transfection assay was performed to examine the function of the β-MyHC gene promoter DHS site. As a positive control, pSV40 CAT (chloramphenicol acetyltransferase) containing the SV40 promoter and enhancer (Promega) was transfected into cardiomyocytes or cardiofibroblasts, and the representative CAT assay is shown in Fig. 20C. As shown in Fig. 20B and C, while the positive control showed similar CAT activity (% of incorporation, raw data see Appendix) in both cell types, a significantly higher CAT activity was observed when the β-promoter/CAT construct was transfected into cardiomyocytes than that of transfection into the cardiofibroblasts. This data suggests that the promoter DHS site is able to drive high level of tissue-specific expression of the chimeric CAT construct. Consistent with our data, Thompson et al. (1991) demonstrated the tissue-specific expression of the cardiac β-MyHC gene in a transient transfection analysis, using a series deletion of rat cardiac β-MyHC gene promoter/CAT chimeric
constructs. In that study, the activity of the CAT reporter gene (-299 upstream of transcription-initiation site) in cardiomyocytes is significantly higher than that of the HeLa cells and 3T3 cells.

12. The 8-2.3 kb site is not a classical enhancer in transient transfection assay

To further investigate the function of the 8-2.3 kb site, two chimeric CAT constructs linking this DHS site (-2477 to -2245) in both orientations to the TK (thymidine kinase) promoter (containing the basal promoter of thymidine kinase gene) were constructed, and then transfected into cardiomyocytes or C2C12 myotubes (Blau et al., 1983). The mouse C2C12 cell line has been shown to express MyoD1, myogenin, and B-MyHC gene (Montarras et al., 1989; Thompson et al., 1991), when differentiated into myotubes. Thus, if the 8-2.3 kb site has enhancer activity, the basal promoter activity of the TK gene should be significantly increased in cardiomyocytes and C2C12 myotubes. As shown in Fig. 21, comparing to the control TK promoter CAT reporter construct, the 8-2.3 kb site linking to the TK promoter in a positive orientation (E+) resulted in 2- to 3-fold (p < 0.01) increase in CAT expression within C2C12 myotubes (Fig. 21). However, no such substantial increase in CAT expression was observed in cultured neonatal cardiomyocytes (see Discussion).

13. GATA-4 activates B-MyHC gene proximal promoter in cardiomyoblasts

In the present experiment and the other study (Thompson et
Fig. 20 **The promoter DHS site is able to confer tissue-specific expression of the chimeric CAT construct** (A) Probe A (see Fig. 7) was used in this study. Lane 1, genomic DNA (20 μg) from late-gestational fetal heart nuclei treated with 2 μg/ml DNase I was digested with NcoI and run in parallel with DNA markers (lane 2). The two DHS sites (8-2.3 kb site and P) are indicated by arrowheads. The upper and lower markers were generated by digesting genomic DNA with NcoI + PstI (-294) and NcoI + XbaI (+97), respectively.
Fig. 20 (continued) (B) Transient transfection analysis of the hamster cardiac β-MyHC gene promoter DHS site/CAT construct (β-294 CAT). Myo, rat neonatal cardiomyocytes; Fib, cardiofibroblasts. The mean ± S.D. of three transfections was plotted. The pRSV β-galactosidase was co-transfected as a control plasmid for transfection efficiency. The percent incorporation of acetylated form in each transfection had been normalized for β-galactosidase activity. (C) A representative CAT assay is shown. Myo: cardiomyocytes; Fib: fibroblasts. In lanes 1 and 3, cells were transfected with pSV40 CAT; in lanes 2 and 4, cells were transfected with β-294 CAT construct.
**Fig. 21** Relative CAT activity of the 8-2.3 kb site in Thymidine kinase (TK) promoter construct. Plasmids were transfected into C2C12 myotubes, cardiomyocytes (Myo) or cardiofibroblasts (Fib). The mean ± S.D. of three transfections was plotted. The percent incorporation of acetylated form in each transfection was corrected for pRSV-β-gal activity before data processing (see Appendix). E(+), positive orientation; E(-), negative orientation (insertion of the 8-2.3 kb site in TK CAT).
al., 1991), low activity of the B-MyHC gene promoter was found in cardiomyocytes, and this gene is inactive in this cell-type. If ectopic expression of the B-MyHC gene can be achieved through an "open" chromatin structure (i.e. with the formation of DHS site), then in transient transfection assay, the introducing of the CAT reporter constructs (in which no chromatin structure is involved) into fibroblasts should result in CAT expression. However, as shown in Fig. 20, when the B-MyHC gene proximal promoter construct (-294 to +97) was transfected into fibroblasts, only limited CAT activity can be detected, suggesting that cardiac-specific nuclear factors are also required in regulating the B-MyHC gene expression. Recently, GATA-4 has been implicated as one of the transcription factors that turn on cardiac-specific gene expression. In the hamster B-MyHC gene promoter, a consensus GATA motif (-271 to -266) juxtaposing the M-CAT motif was identified, and mutation of this GATA-DNA binding site resulted in decreased CAT activity by 60% in cardiomyocytes (Fig. 22A). This reduction in CAT activity correlated with a loss of DNA-binding activity when the mutant GATA motif was used as a probe in a gel mobility shift assay (Fig. 22B lane 7). However, when the B-MyHC gene promoter GATA motif (-279 to -254) was used as a probe, GATA-DNA binding activity was detected in lane 2 (Fig. 22B, arrowhead). The DNA-binding specificity was established by the observation that the shifted band could be competed away by the GATA-2 binding site of the PPET-1 gene (lane 4) or by a large excess of cold probe (lane 3), but competed less effectively by a mutant B-MyHC gene GATA motif (lane
6, nt sequence see Materials and Methods), or an oligomer containing GC rich residues (CT/ACCC box, lane 5).

When the GATA-2 binding motif was used as a competitor, the upper shifted band (arrowhead) was almost completely competed away (Fig. 22B, lane 4). However, when the specific cold probe was used as a competitor, in addition to the upper shifted band, the three lower shifted bands (B1, B2, B3) were also eliminated (lane 3). This discrepancy in competition suggests that there exists multiple GATA-related factors (see Discussion) in the cardiac nuclear extracts. One of the GATA factors may have the same binding specificity between the GATA-2 binding site and the GATA motif of the B-MyHC gene, while others preferentially bind to the B-MyHC gene GATA motif.

To further establish the role of GATA-binding factors in the regulation of B-MyHC gene expression, the GATA-4 expression vector was co-transfected with the B-MyHC promoter/CAT construct into fibroblasts. The forced expression of GATA-4 in fibroblasts lead to marked transactivating of the B-MyHC promoter (5- to 8-fold). No such enhancing activity can be observed when GATA-4 was co-transfected with the mutant B-MyHC promoter (Fig. 22C).
Fig. 22 Transactivation of the β-MyHC gene promoter by GATA-4 (A).

The wild-type β-294 promoter/CAT construct or the mutant β-294 promoter/CAT construct (mβ-294) was transfected into cultured neonatal cardiomyocytes. The results were plotted according to three independent experiments, and expressed as percent incorporation of acetylated form after normalized for pRSV β-gal activity. T test indicated significant difference between β-294 and mβ-294 (p< 0.01).
Fig. 22 (continued) (B) Gel mobility shift analysis using the β-MyHC gene promoter GATA motif (β-GATA, lanes 2-7) or mutant GATA motif (mu, lane 7) as a probe in the absence (lane 1) or presence (lanes 2-7) of cardiomyocytes nuclear extracts. The various competitors used are as indicated. The abbreviations used are: C, CT/ACCC box; S2, PPET-1 gene GATA-2-binding site; Sp, cold probe; Cp, competitor. In lane 6, the competitor is the cold mutant probe. The three lower shifted bands in lane 2 are indicated by B1, B2, and B3, respectively.
Fig. 22 (continued) (c) the pMT-2 GATA-4 expression vector (10 µg) was co-transfected with the B-MyHC promoter CAT construct (15 µg, 8-294) or the mutant B-MyHC promoter/CAT plasmid (m8-294) into fibroblasts. As control, the same co-transfection experiments were repeated, using the pMT-2 vector instead of the pMT-2 GATA-4 vector. The data shown are the means of three independent experiments. The percent incorporation of acetylated form in each transfection had been corrected for β-gal activity, and the results are expressed as fold activation relative to transfection with B-MyHC promoter CAT construct.
(V) Molecular cloning of human GATA-4 cDNAs

14. Isolation and characterisation of the human GATA-4 cDNA clones

Based on previous studies that demonstrated the role of GATA-4 in the regulation of the cardiac MyHC gene expression (Molkentin, 1994), we then focused on characterizing the human GATA-4.

A 246 bp cDNA sequence that encodes the conserved zinc finger DNA-binding domain of the human GATA-4 was generated by RT-PCR, using a pair of oligo primers derived from murine GATA-4 amino acid sequences (see Materials and Methods). This 246 bp DNA fragment was then used as probe to screen a human fetal heart cDNA library. Four cDNA clones hybridizing to the probe were obtained by screening approximately $1 \times 10^7$ recombinant phage plaques. Two of the clones (HGA-1 and HGA-14) overlapped and contained a long open reading-frame encoding the hGATA-4 (The GenBank accession number is L34357). The cDNA and predicted amino acid sequences for hGATA-4 are shown in Fig 23. The hGATA-4 cDNA sequence within the protein coding region shows 85% identity with murine GATA-4.

Two potential translation initiation sites were found in the sequence of hGATA-4. The predicted molecular weight of the hGATA-4 polypeptide based on these two putative start sites is 48kDa. However the nucleotide (nt) sequence(GCCAUGG) flanking the second Met codon at nt 259 is identical to the reported optimal sequence for eukaryotic initiation of translation[(A/G)CCAUGG] (Kozak, 1986). This Met codon is proposed to be the putative translation initiation site of the hGATA-4.
15. Northern blot analysis of the hGATA-4

As shown in Fig 24, the hGTAT-4 is expressed both in fetal (lane 2) and adult heart ventricle (lane 1), and is not detectable in brain (lane 4) or liver (lane 5) as revealed by Northern blot analysis. Since GATA-4 of small rodents is expressed in heart, small intestine and gonad (Arceci et al., 1993; Grepin et al., 1994), rat ovary is also included in the same blot for comparison. The transcript size of hGATA-4 is similar to the rat; both are 4.4 kb.
Fig. 23 The nucleotide sequence of the human GATA-4 (hGATA-4) cDNA and deduced amino acid sequence. The first ATG and two zinc-finger motifs (DNA-binding domains) are underlined.
Fig. 24 Northern blot analysis of hGATA-4. The tissues used were human adult heart (lane 1) and fetal heart (lane 2), brain (lane 4) and liver (lane 5), and rat ovary (lane 3). The positions of 28S and 18S of rRNA are indicated. The ethidium-bromide-stained gel is also shown.
DISCUSSION
Chromatin remodelling of the cardiac B-MyHC gene

Transcriptional regulation of the cardiac B-MyHC gene has been demonstrated by Northern blot (Lompre et al., 1984), CAT assay (Thompson et al., 1991; Flink et al., 1992; Schimizu et al., 1992; Cribb et al., 1989) and nuclear run-on analysis (Boheler et al., 1992). In the present study, we extend those earlier investigations, and provide the first evidence that the isoformic switch of cardiac B-MyHC gene during cardiac development is accompanied by an alteration of the nucleosomal structure of its regulatory elements. Such changes in DHS patterns during cardiogenesis strongly suggests a role for chromatin remodelling in the regulation of cardiac-specific gene expression.

It would be informative to compare our data with Northern blot analysis of the cardiac MyHC genes. These two assays (i.e., Southern blot analysis for DNase I hypersensitive sites and Northern blot analysis) are fundamentally different in sensitivity. Northern blot analysis is useful for detecting major transcripts (mRNA) in tissues, and the Southern blot analysis is used to detect single copy genes in the genome. For example, using Southern blot analysis, we showed that the B-promoter DHS site became barely detectable in 3-day-old heart (Fig. 9B lanes 7-8). In Northern blot analysis, low level expression of the B-MyHC transcript remained observed in 10-day-old heart (Fig. 27B, lane 2). Moreover, the transcripts detected by Northern blot cannot necessarily be taken to provide evidence of gene activity because this analysis reflects the steady level of mRNA (the balance of the rate of production and
the rate of degradation). Therefore, even the rate of transcription decreased and the intensity of the β-promoter DHS site diminished (Fig. 9B lanes 4-6), abundant β-MyHC mRNA (because of its stability) still can be detected by Northern blot analysis in one- to two-day-old heart (Fig. 27B, lane 1). Comparing the data between these two assays in an age-matched manner would be inappropriate.

Our data (the disappearance of the β-MyHC gene promoter DHS site in adult hamster, as shown in Fig. 9), however, correlates well with the general expression pattern of the cardiac β-MyHC gene observed during cardiac development of small mammals. Many investigators have shown that the β-MyHC gene is active in the fetal heart ventricles and inactive in the adult hearts of small mammals (Lompre et al., 1984; Lyons et al., 1990; Morkin, 1993; Wang, 1994). The activity diminishes gradually: for example, in the mouse, the β-MyHC gene is down-regulated just after birth, but β-MyHC transcript persists for about 7 days after that time and is then replaced by the α-MyHC transcript in the ventricular muscle (Lyons et al., 1990). These data and our results suggest that, in the early postnatal stage, the heart may be constituted by a heterogeneous population of cardiomyocytes that are at different points in the maturation pathway (Lompre et al., 1984). Thus, it is possible that the decreased intensity of the β-MyHC gene promoter DHS site in neonatal cardiomyocytes (as compared to late-gestation fetal cardiomyocytes) is due to maturation lags in different cardiomyocyte subgroups. In some cells, the β-MyHC promoter remains DNase I hypersensitive; in some, however, the β-
MyHC gene promoter is no longer DNase I hypersensitive.

The chromatin structure of the cardiac β-MyHC gene was also probed by MNase. As shown in Fig. 11, the usual nucleosomal repeat in the transcribed region becomes less prominent in fetal hearts, indicating that nucleosomal organization has changed. This result is consistent with the observation that the expected nucleosomal repeat of some heat shock genes is replaced by a continuous smear when intense transcription begins (Elgin, 1988).

The nature of this change remains unknown. The usual nucleosomal repeat may disappear because it is obscured by the presence of RNA polymerase II, because the nucleosomes have been modified or displaced, or because there is a disruption of higher-order structure.

By contrast, the 5′ proximal promoter region of the β-MyHC gene is organized in an array of specific-positioned nucleosomes when it is actively transcribed. Interestingly, the promoter DHS site is mapped to one of the nucleosomes (N2 particle), implying that the N2 particle is perturbed. Many transcription factor-binding sites crucial for the regulation of this gene (see Fig. 4 and 11D) happen to be localized in this nucleosome. The nature of this perturbed nucleosome remains unknown. Although its basic structure (DNA wound around an octameric histone core) may remain intact, the histones may be modified by acetylation or by other mechanisms leading to DNase I hyper-sensitivity.

The above results are reproducible. Both liver and fetal heart nuclei used in Fig. 11B and C were isolated from two independent
preparations. The general nucleosomal structure of liver and fetal heart chromatin observed in Fig. 11B (lanes 1-4 for liver, and lanes 5-7 for fetal heart) can be repeated in Fig. 11C (lanes 4-6 for liver, and lanes 7-9 for fetal heart). However, in both figures, higher background and extra bands were also observed, which might be caused by the promiscuous nature of MNase which is able to digest linker DNA between individual nucleosomes and to attack AT-rich sequences in the genome (McPherson et al., 1993).

Limitations of DHS sites approach

There are at present at least three potential limitations to the DHS sites approach described. First, in order to prove the functions of the DHS sites, a sophisticated biotechnology (e.g., transgenic study) is often required. Second, the only genes that can be studied are those for which a detailed restriction enzyme map is available. Third, large amounts of nuclei are required for testing for the optimal DNase I concentration for digestion. Thus, if tissue availability is limited (as is the case with human tissues), it becomes difficult to perform this experiment. Most early work in identification of DHS was carried out using nuclei from Drosophila, yeast, or tissue cell lines. As well, because liver nuclei are abundant, the DHS of the albumin gene have been studied extensively (Liu et al., 1988).

It would have been much more convincing to demonstrate a DHS site in the presence of a very low concentration of DNase I. The lack of such a demonstration in the present study or an earlier
study (Liu et al., 1988) may be due to impurities of a nuclei preparation from whole tissues. Since the cardiac nuclei are prone to be trapped in the abundant connective tissues during nuclei isolation, the purity and yield rate is frequently unsatisfactory in the cardiac nuclei preparation (Jackowski and Liew, 1980). Therefore, the cardiac nuclei had to be prepared constantly throughout the course of the present study. Fluctuation of nuclei purity becomes inevitable. In general, the more initial tissue used, the lower the purity. The purity of nuclei was reflected by the amount of DNase I required to detect the DHS sites. It has been suggested that higher concentration of DNase I is required to probe the chromatin structure of crude nuclei preparations (Wu 1989). Although time-consuming to perform repeated cardiac nuclei preparation, we performed these studies and demonstrated the reproducibility of our findings.

DNase I sensitive domain

Chromatin displays two types of DNase I sensitivity. The first, DNase I hypersensitivity, is localized to short, 200 bp stretches of chromatin (Elgin, 1988), while the second, domain DNase I sensitivity (Stalder et al., 1980a; Weintraub and Groudine, 1976; Wood et al., 1982), extends over 12 to 100 kbs of chromatin and includes both nontranscribed and transcribed DNA sequences (Alevy et al, 1984; Jantzen et al., 1986; also see Introduction p. 23). An increase in domain DNase I sensitivity has been reported to be an early event in the activation of gene expression (Stalder
et al., 1980; Feng and Villeponteau, 1990). For example, within 90 seconds of serum induction, the c-fos proximal DNA sequences on both sides of the 5' enhancer exhibit increased DNase I sensitivity. Within 5 min, elevated DNase I sensitivity spreads to chromatin at the distal 3' end of the c-fos gene. Furthermore, the induced changes in chromatin structure precede the increased transcriptional activity of the c-fos gene. (Feng and Villeponteau, 1990).

Because of the DNase I sensitive nature of the active chromatin domain, a one- to two-fold increase in DNase I concentration may lead to a significant decrease in the intensity of both major band and DHS sites of the β-MyHC gene (during which stage the gene is active). DNase I sensitive activity is easier to observe when the gene is quite active and the nuclei preparation is pure (for example, in Fig. 8B, lanes 1 and 2, and Fig 9B lanes 1 to 3, considering that the β-MyHC gene is much more active in the fetal heart than that of the neonatal heart). Our results are consistent with DHS site data reported for the myogenin gene (Gerber et al., 1997). In that study, a DHS site of weak intensity was first detected at DNase I concentrations of 0.2 unit; and intensity was strongest in the presence of 0.6 unit of DNase I; when DNase I concentration was doubled (1.2 unit), the intensity of both the major band and the DHS sites decreased significantly.

Thus, the cardiac DHS sites detected in the present study may be defined as DNase I hypersensitive for the following reason: although the entire gene is sensitive to DNase I, some regions of
the gene can clearly be seen to be more sensitive (hypersensitive), i.e., they form distinct darker subbands against the background, as shown in Southern blot analysis. These subbands are also tissue-specific (i.e., we were unable to detect these sites in other tissues, no matter what concentration of DNase I was used).

Cardiac Hypertrophy

In response to hormonal, physiological, hemodynamic, and pathological stimuli, adult ventricular muscle cells can adapt to an increased workload by activating a hypertrophic process, characterized by an increase in the contractile protein content of individual cardiac muscle cells with a concomitant activation of embryonic genetic program (for review see Chien et al., 1993). Although hypertrophy would seem to be a benefical adaptation to alterations in hemodynamic loading, it is also clear that this growth process is frequently maladaptive (Schwartz et al., 1993; LeJemtel and Sonnenblick, 1993). For example, the hypertrophied myocardium in cardiac diseases appears to exhibit frequently abnormal contractility and electrical instability (Gwathmey and Morgan, 1985; LeJemtel and Sonnenblick, 1993). Thus, an improved understanding of the mechanisms that lead to the reactivation of the embryonic genetic program during cardiac hypertrophy, will ultimately lead to the identification of the signalling pathways that cause this adaptive/maladaptive response.

The reactivation of β-MyHC gene expression is one of the first well-characterized molecular markers in the hypertrophic heart.
Studies of pressure overload hypertrophy in rat model systems have documented a qualitative change in myosin composition from the V1 to the V3 isoform (see Introduction, p.30; Nagai et al., 1987). It is now known that the alteration in myosin composition is the result of the down-regulation of the α-MyHC gene and the concomitant reactivation of β-MyHC gene expression, which is normally expressed in embryonic ventricular muscle cells (Izumo et al., 1987). Accordingly, the mechanisms that cause such a switch in gene expression have been under intensive investigation during the past decade.

Cardiac β-MyHC gene induction in the hypertrophic response has been shown to be mediated by a number of agonists (external signalling molecules), such as α adrenergic receptor agonists (Kariya et al., 1991), norepinephrine (Chen et al., 1992), endothelin-I (Wang et al., 1992), and stretch stimulation (Sadoshima et al., 1992; Shyu et al., 1995). In those studies, the cis-acting elements mediating the induction of β-MyHC gene expression were identified by transiently transfecting the β-MyHC promoter/CAT reporter constructs into cultured cardiomyocytes. For example, recent studies by Simpson and co-workers have localized within the rat β-MyHC promoter region the M-CAT motif that mediates inducible expression during α-adrenergic stimulation of myocardial cell hypertrophy (Kariya et al., 1993 and 1994). Sadoshima et al., (1992) however, found that the stretch-response element of the rat β-MyHC gene lies beyond -628 bp of the 5′ flanking region. The norepinephrine response element was mapped to yet another region of
the β-MyHC promoter (Chen et al., 1997). So far no consensus hypertrophic-response element(s) have been localized within the β-MyHC gene promoter, prompting the question as to whether the in vitro transient transfection system is an appropriate model for investigating the induction of the β-MyHC gene in vivo.

The β-MyHC promoter consists of an array of clustered cis-acting elements: the thyroid hormone response elements, Bel, Be2 (M-CAT), GATA motif (this study), C-rich regions, Be3, Be4, Be5, and Be6 (Thompson et al., 1991). Since gene regulation is different between transiently transfected promoter constructs and promoters organized in a nucleosomal structure (see Introduction, "Naked DNA versus Nucleosomal DNA"), we hypothesize that a unique nucleosomal structure is required for those nuclear factors to bind to their binding sites simultaneously. Therefore, it is proposed to use the β-MyHC promoter organized into the nucleosome structure as a model to study the reexpression of the β-MyHC gene in pathological conditions.

Chromatin remodelling and cardiac hypertrophy

Perturbation of nucleosomal structure (formation of the DHS site) appears to be an initial step in gene activation (Elgin, 1988). It is predicted that reexpression of the β-MyHC gene in cardiac hypertrophy will be accompanied by the reappearance of the β-MyHC promoter DHS site. Therefore, it would be interesting to investigate the signaling pathways involved in the reprogramming of chromatin structure to allow transcription factor binding in the
diseased heart.

The reappearance of the B-MyHC promoter DHS site in the hypertrophic heart could be tested by the indirect-end labelling method. However, this methodology may be limited in its sensitivity, especially if the cardiac B-MyHC gene is induced only in a subpopulation of cardiomyocytes. In fact, in one study, the induction of the cardiac B-MyHC gene was found to be localized to perivascular areas (Schiaffino et al., 1989). Thus, even though the cardiac B-MyHC mRNA is significantly increased in Northern blot analysis (a sensitive method in detecting major transcripts in a tissue), the reappearance of DHS site may not be detected by the indirect end-labelling method (which is based on Southern blot analysis of a single copy gene in the genome).

A direct link between hypertrophic stimulation and chromatin structure alterations in cardiomyocytes has been demonstrated by Chen and Liew (1993). In that study, a highly phosphorylated 31kd protein was identified in response to the administration of norepinephrine to cultured cardiomyocytes. The increased phosphorylation of this protein was observed for 5 min, and lasted at least for 2 hrs (the maximum stimulation time of this study) after the administration of norepinephrine. This effect is specific to norepinephrine, since endothelin-1, TPA, or platelet activating factor were unable to induce a similar response. Furthermore, treating the cells with terazosin, an α1-adrenergic receptor blocker, 15 min before administration of norepinephrine, inhibited the phosphorylation. Peptide-microsequencing of this protein
revealed that it is one of the subtypes of H1 (Chen and Liew, unpublished data). This result is interesting, since H1 in active chromatin is always depleted, and phosphorylation of H1 has been shown to reduce its ability to condense chromatin (Felsenfeld, 1992). Whether other histone modifications occur in the hypertrophic heart could be examined in future experiments.

**DHS Sites and Familial Hypertrophic Cardiomyopathy**

In the present study, the mapping of the cardiac MyHC gene DHS sites was focusing on the 5' upstream region of the cardiac MyHC gene. In future work, it would be interesting to map the DHS sites downstream of the transcription initiation sites of the cardiac MyHC genes. For example, some of the DHS sites found in the downstream regions of the β-MyHC gene may contain the crucial cis-acting elements regulating the expression of the cardiac α-MyHC gene. Some DHS sites may be associated with other important roles for gene functions, such as recombination or chromosome segregation (Gross and Garrard 1988).

An α/β cardiac MyHC hybrid gene has been implicated as one of the mutations causing familial hypertrophic cardiomyopathy (FHC) (Tanigawa et al., 1990). This hybrid gene has a unique structure, in that its 5' end and 3' end are derived from the α-MyHC gene and β-MyHC gene, respectively. In the hybrid gene, the transition between the α- and β-MyHC gene occurs at the exon 27 of the β-MyHC gene. It has been suggested that the hybrid gene is generated by an unequal recombination event occurring between the α- and β-MyHC
genes. In this regard, it would be of particular interest to examine the DHS sites around exon 27 of the β-MyHC gene. Since the exons of the cardiac MyHC genes are highly conserved across species (Wang 1994), the results obtained from the hamster cardiac MyHC gene may be extended to the human MyHC genes.

Two mechanisms by which mutant myosin proteins may cause defective muscle formation have been documented (Bejsovec and Anderson, 1988; Beall et al., 1989; Elstein et al., 1992). (1) Mutations cause the defective muscle protein unable to bind to other components of the sarcomere (Bejsovec and Anderson, 1988). (2) Altering the ratio of MyHC to other polypeptides leads to aberrant formation of muscle sarcomere (Beall et al., 1989). Thus, a mutation affecting the transcription rate of the β-MyHC gene may also lead to the diseased phenotype. In this regard, the 5' DHS sites of the β-MyHC gene would be the candidate regions associated with mutations. So far, no mutation occurring at the 5' upstream region of the cardiac β-MyHC gene has been reported.

The role of thyroid hormone and its receptor in the formation of DHS sites

To elucidate the sequence of events that results in developmentally regulated changes in chromatin structure remains a challenging problem. It has been shown that a circulating thyroid hormone surge, which occurs at about the time of birth, causes both up-regulation of the cardiac α-MyHC gene and concomitant down-regulation of the cardiac β-MyHC gene (Morkin et al., 1993).
Thyroid hormone can also induce this switch in gene expression in cell culture with serum-free medium (Gustafson et al., 1987). Studies at the protein level showed that injecting thyroid hormone into timed pregnant rats (Chizzonite and Zak, 1984), caused the isoformic switch of MyHC proteins in fetal heart.

In the present study, thyroid hormone was found to have an effect on chromatin structure. It was found that the administration of thyroid hormone causes the switch of fetal DHS sites patterns to the adult patterns (Fig. 10).

How thyroid hormone causes chromatin remodelling of the β-MyHC gene remains unknown. However, several lines of evidence do indicate that its nuclear receptor, thyroid hormone receptor (TR), may have effect on chromatin structure. For example, using a yeast two-hybrid system, Lee et al. (1995) found that the thyroid hormone receptor β (TRβ) can, when coupling with its ligand, interact with HMG-14 and HMG-17 (HMG: high mobility group proteins) which contact the DNA near the entry/exit point of the histone core (Alfonso et al., 1994) and are thought to be associated with active chromatin. In addition, recently, Wong et al. (1995), using Xenopus TRA gene (thyroid hormone receptor A gene) as a model, demonstrated that TR coupling with its ligand can direct the disruption of local chromatin structure. Since TR and GR are within the same nuclear receptor superfamily (Izumo and Mahdavi, 1988; for review see Beato et al., 1995), it has been suggested that the TR-induced nucleosomal change may be related to the recruitment of SWI/SNF complex or of other chromatin remodelling factors (Wong et al.,
The most compelling evidence that directly links TR to chromatin remodelling comes from the discovery that its coactivators, p300/CBP, are histone acetyltransferases (Ogryzko et al., 1996; Bannister and Kouzarides, 1996), and its corepressor, N-CoR (nuclear receptor corepressor) is one of the components of a histone deacetylase complex (Heinzel et al., 1997; Alland et al., 1997; Pazin and Kadonaga, 1997). With those discoveries, an attractive model was proposed (Brownell et al., 1996; Wolffe, 1997). In that model, TR can modify the chromatin structure of its target genes through the recruitment of histone acetyltransferase (in a TH-dependent manner) or deacetylase (without TH) to the promoter regions. Thus, hyperacetylation allows basal transcription factors to bind to the promoter regions, leading to the initiation of gene expression; by contrast, deacetylation precludes the transcription factor binding and repress gene expression. However, this model can only be applied to explain the genes whose activation is TH-dependent, but not to the genes (such as the Β-MyHC gene), which become inactivated upon treatment of TH.

The -2.4 kb silencer in the chicken lysozyme gene provides another interesting example of the interplay between thyroid hormone negative regulation and the formation of DHS sites (Banaihamad et al., 1990). This silencer was first identified in a DNase I hypersensitive site study (Fritton et al., 1984). This DHS site is located -2.4 kb upstream of the transcription start site of the chicken lysozyme gene. The site is present in oviduct,
liver, kidney, brain and in 5-day-old embryos, but is absent in macrophages and erythrocytes (Fritton et al., 1984). Two nuclear factor binding sites were identified within this site by footprinting analysis, and one of the binding sites is a thyroid hormone receptor binding site. In functional analysis, this TRE (thyroid hormone response element) can exert repression or thyroid hormone induction on several promoter CAT constructs. How thyroid hormone and TR negatively regulate gene expression in the context of chromatin structure was not addressed in that study.

The integrity of cardiomyocyte nuclei

Rather than using a cell line, mammalian cardiomyocyte nuclei were first used in the present study to investigate the chromatin structure of cardiac-specific genes. The integrity of nuclear preparation was demonstrated by the observation of the expected nucleosome repeat after MNase digestion (Fig. 11A). Furthermore, chromatin remodelling of the cardiac MyHC genes was demonstrated by comparing cardiomyocytes at various developmental stages. This demonstration suggests that the method of preparation of heart ventricular nuclei developed in our laboratory (Jackowski and Liew, 1980) is suitable for the study of chromatin structure of cardiac-specific genes. In fact, ventricular nuclei preparation by this method could also be used in nuclear run-on transcription assays (Boheler et al., 1992).

The 8-2.3 kb site
The identification of the 8-2.3 kb site is consistent with the findings of transgenic studies (Rindt et al., 1993; Knotts et al., 1994, also see Introduction p.37), which detected an undefined cis-acting element located within the 5 kb upstream region of the β-MyHC gene. Whether there are other cis-acting elements residing further upstream remains to be determined. In future experiment, the position of the 8-2.3 kb site can be more precisely localized by using a second probe whose position in the β-MyHC gene is closer to the 8-2.3 kb site (Wu, 1989).

Significantly, a MEF-2 binding site was identified within the 8-2.3 kb site. The MEF proteins, members of the MADS box family that includes serum response factor (SRF, p67SRF), are involved in the regulation of muscle gene expression (for review see Shore and Sharrocks, 1995). Loss-of-function mutations in the single mef2 gene in Drosophila prevent the differentiation of somatic, cardiac, and visceral muscle cells, although the specification and early development of these tissues appears unaffected (Bour et al., 1995; Lily et al., 1995; Ranganayakulu, 1995). In fact, in mutant embryos, the cardiac cells do not express MyHC (Bour et al., 1995). These results indicate that MEF-2 is required for later aspects of the differentiation of the three major types of musculature.

In the present study, the biological function of the 8-2.3 kb site was examined by CAT assay analysis. It was expected that when the 8-2.3 kb site was linked to the TK promoter, a significant increase in CAT activity in cardiomyocytes would result. However, as shown in Fig. 21, no substantial increase in CAT activity was
observed. The lack of contrast between these two constructs may be caused by our inappropriate CAT assay conditions, in which the amount of proteins (CAT) used in reactions is in a large excess of the optimal concentrations required for CAT assays. Such a possibility can be tested by using less amount of protein for CAT assays in future experiments.

Alternatively, the lack of activity might be because (1) cultured neonatal cardiomyocytes may not be the physiological environment for the expression of β-MyHC gene, as the β-MyHC gene is down-regulated right after birth, (2) crucial nuclear factors that are required for the β-2.3 kb site to function properly are not expressed in neonatal cardiomyocytes, or (3) the cooperative interaction of multiple transcription factors binding to the β-2.3 kb site requires a precisely positioned nucleosome over this region (McPherson et al., 1993), (in this site, the distance between the first E-box and MEF-2 binding site is 160 bp, which is the length of two full turns of DNA wrapping around a nucleosome).

The β-2.3 kb site appears not to act as a typical enhancer (which is supposed to be active independent of the orientation in the reporter construct) in transient transfection assay, despite the fact that this site is associated with a cluster of conserved muscle regulatory motifs. One of the possible roles of the β-2.3 kb site is that it functions in the establishment, maintenance, or both, of an open chromatin domain that confers the position-independent and high level transgene expression. Such a possibility can be tested in a transgenic study, or by using the homologous
recombination technology to delete this DHS site in mouse embryonic stem cells (Fiering et al., 1995; Hug et al., 1996)

The α-1.9 kb site

The role of the α-1.9 kb site in vivo remains to be determined. Transgenic analysis (see Introduction, p.33; Subramaniam, 1993) has indicated that an undefined regulatory element is located in the distal 2.8 kb segment of the 4 kb upstream region of the cardiac α-MyHC gene (i.e., the intergenic region between the β- and α-MyHC genes). Since the α-1.9 kb site is located in this distal 2.8 kb segment, this DHS site might be associated with the candidate regulatory element. However, so far no upstream regulatory elements have been identified in transient transfection assays; therefore, a more detailed genetic assessment of this site in the context of chromatin structure is required.

The biological significance of the α-1.9 kb site is demonstrated by its association with a conserved GATA motif and provides further evidence that GATA-4 plays a role in the regulation of cardiac α-MyHC gene expression (Molkentin et al., 1994; Huang et al., 1995). However, it should be noted that the α-1.9 kb site is also located at the 3’ end of the cardiac β-MyHC gene and regulatory elements (such as an enhancer) located at the 3’ end of a gene has been documented before (Wall et al., 1988). Whether the conserved GATA motif within the α-1.9 kb site also plays a role in the regulation of the β-MyHC gene remains to be tested.

125
GATA-4 and chromatin remodelling

A functional GATA motif within the β-MyHC gene proximal promoter region was identified (Fig 22A). Furthermore, GATA-4 could activate β-MyHC gene promoter/CAT construct in fibroblasts (Fig. 22C). These results, together with the observation that inhibition of GATA-4 expression causes the down-regulation of the β-MyHC gene expression in an in vitro cardiac muscle differentiation system (Grepin et al., 1995), strongly indicate that GATA-4 plays a role in the regulation of the cardiac β-MyHC gene expression.

Several additional cardiac genes have been shown to be regulated by GATA-4, including cardiac troponin C (Ip et al., 1994), cardiac α-MyHC (Molkentin et al., 1994), and atrial and brain type natriuretic peptides (Grepin et al., 1994). The cardiac troponin C promoter is transactivated by transient expression of GATA-4 in fibroblasts (Ip et al., 1994). Although the promoter regions of the rat, rabbit and human β-MyHC genes have been characterized (Thompson et al., 1991; Cribbs et al., 1989; Flink et al., 1992), and a number of potential regulatory elements within 300 bp of the transcriptional start site have been identified, the present study provides evidence that GATA-4 also regulates the β-MyHC gene expression.

The GATA factors are a family of transcriptional regulators that are expressed in a tissue-restricted manner (for review see Orkin, 1992). GATA-4 mRNA is most abundant in heart, with low level in gonads and small intestine (Arceci et al., 1993; Huang et al.,
1995). In the mouse embryo, at 3 days postcoitum, GATA-4 mRNA is expressed in endocardium, myocardium, and precardiac mesoderm. By the onset of cardiac septation (9 days postcoitum), GATA-4 mRNA expression is evident in endocardium, endocardial cushion tissue, and myocardium (Heikinheimo et al., 1994). Expression of GATA-4 mRNA in the myocardium continues throughout gestation and into the postnatal period (Heikinheimo et al., 1994; Huang et al., 1995). The other two members of this family, GATA-5, and -6, were isolated in a chicken embryo cDNA library by low-stringency screening (Laverriere et al., 1994). The temporal and spatial patterns of GATA-4, -5, and -6 expression support a role for these factors in the regulation of cardiac differentiation, analogous to the established role of transcription factor GATA-1 in the regulation of haematopoiesis (Tsai et al., 1989; Pevny et al., 1991).

As demonstrated in Fig. 24, GATA-4 mRNA is expressed continuously from the early fetal to the adult heart stages. One question that remains however, is how postnatal down-regulation of the β-MyHC can occur in the presence of GATA-4. Based on the disappearance of the promoter DHS site in adult heart nuclei, it is hypothesized that the stage-specific expression of the β-MyHC promoter is achieved by the formation of a "closed" chromatin structure that does not allow GATA-4 to bind to its promoter in adult heart. Alternatively, GATA-4 can bind to the β-MyHC promoter compacted in nucleosomal structure, but is unable in such a "closed" chromatin structure, to activate gene expression. This hypothesis also predicts that GATA-4 does not possess the ability
to bind to or to disrupt the organized nucleosome structure in promoter regions (at least in the β-MyHC gene promoter). Such an assumption is consistent with Armstrong and Emerson's observation (1996) that GATA-1 (which contains a highly homologous DNA-binding domain with GATA-4 (Arceci et al., 1993)), is unable to bind to its binding site over nucleosomal DNA. In that study, the β-globin gene DNase I hypersensitive site 2 (HS-2) was reconstituted into an array of nucleosomes, using an in vitro chromatin assembly system developed by Kadonaga and his co-workers (Kamakaka, et al., 1993). Only after this chromatin-reconstituted HS-2 was disrupted by binding to another erythroid-specific nuclear factor (NF-E2), was GATA-1 allowed to bind to its binding site.

Since the pattern of developmentally-specific ANP gene expression is very similar to that of the cardiac β-MyHC gene, it is tempting to apply the above hypothesis to explain the regulation of atrial natriuretic peptide (ANP) gene expression. The ANP mRNA is expressed throughout embryonic and fetal development in atrial and ventricular cells (Zeller et al., 1987). However, in the latter stages of fetal development the ANP gene is switched off in ventricular cells, whereas its level of expression remains high in atria, thus establishing the adult pattern of expression for this gene (Argentin et al., 1994). The decrease in ventricular ANP expression corresponds temporally to the arrest of cell division in the ventricular myocardium, which occurs around the time of birth in mammals (Claycomb, 1975), and the ANP gene is re-expressed in hypertrophied ventricular cells (Chien et al., 1991). As
mentioned previously, a functional GATA motif was identified in the ANP proximal promoter region (Grepin et al., 1994). Thus, adopting a model of chromatin remodelling allows us to explain the complex patterns of some aspects of cardiac-specific gene expression. However, the results demonstrated in this study and the simple model proposed above do not exclude other possible mechanisms (for example, DNA methylation, see Introduction) that might mediate the differential regulation of cardiac gene expression (Flink and Morkin, 1995; Mably and Liew, 1996).

Conclusions

1. These studies provide the first evidence for chromatin remodelling in the developmental regulation of cardiac gene expression. The β-MyHC gene promoter DHS site is mapped to a specific positioned nucleosome only in fetal hearts, suggesting that nucleosomal structure may participate actively in gene regulation.

2. Two far upstream DHS sites of the cardiac MyHC genes (the α-1.9 kb site and β-2.3 kb site) associated with muscle-specific regulatory elements were identified. The novel regulatory elements identified in the far upstream region of the cardiac MyHC genes helps to extend our understanding of the regulation of these genes.

3. These studies demonstrated that thyroid hormone is able to alter the chromatin structure of the cardiac β-MyHC gene, suggesting that this hormone and its nuclear receptor may play a key role
in chromatin remodelling of the MyHC genes.
4. These studies showed that GATA-binding factors also regulate the expression of the cardiac β-MyHC gene.
5. Thus, the appropriate spatial and developmental pattern of cardiac gene expression requires complex regulatory pathways, involving interplay between chromatin structure and tissue- (and stage-) specific nuclear factors.

Future Experiments (Fig. 25)
The cleaved subbands for the β-2.3 kb site and the α-1.9 kb site are of a length 4 kb and 2 kb, respectively. Such crude mapping can be further defined by selecting a new restriction site as a point of reference (and a new probe) such that the cleaved subband is of a length (0.5 to 1.5 kb) that is maximally resolvable on long agarose gels (Wu, 1989).

Whether DNA methylation also plays a role in the cardiac MyHC gene expression can be investigated using the similar approach as the identification of nuclease hypersensitive sites. In this study, methylation-sensitive restriction endonucleases (instead of DNase I) will be used to assess the methylation status of a given CpG dinucleotide in the DHS sites identified previously.

DHS sites located in the other regions of the cardiac MyHC genes will be mapped. It would be interesting to investigate whether allelic variants are associated with these DHS sites. Significant allelic variants detected in a standard sequencing analysis may be one of the underlying causes of familiar hyper-
trophic cardiomyopathy.

In order to dissect the functions of the two distal DHS sites (8-2.3 kb and a-1.9 kb sites), transgenes with or without these sites will be created, and their developmental expression patterns will be compared.

Reconstitution of the cardiac MyHC gene promoter DHS sites with chromatin structure in vitro and in vivo should allow us to examine the relationship between transcription factors and chromatin in the formation of the DHS site. Specifically, by using the in vitro and in vivo chromatin assembly systems described previously (Croston, et al., 1992; Kamakaka, et al., 1993; Almouzni and Wolff, 1993), the role of thyroid hormone receptor, M-CAT factor, GATA-4, and Sp-1 in the regulation of cardiac MyHC genes with chromatin template can be examined biochemically, and the formation of the DHS site (or disruption of their nucleosomal structure) can be revealed by digestion of the chromatin-reconstituted MyHC gene promoters with DNase I and MNase.

Considerable evidence establishes the usefulness of these chromatin assembly systems (Croston, et al., 1992; Pazin, et al., 1996; Kamakaka, et al., 1993; Armstrong and Emerson, 1996; Wong et al., 1995). The reconstituted chromatin in vitro, consists of periodic arrays of fully assembled nucleosomes with properly incorporated histone H1 and physiological nucleosome spacing and is therefore ideal for the study of chromatin structure and transcription. In the mammalian genome, the transcription of the HIV-1 enhancer (Pazin et al., 1996) and the formation of the Human
β-globin locus control region hypersensitive site 2 (HS-2) (Armstrong and Emerson, 1996) have been studied using this system, and provide valuable findings regarding the role of NF-κB and NF-E2 in transcriptional activation of the HIV-1 enhancer and the formation of β-globin HS-2, respectively. As well, by staged injection of mRNA encoding transcription factors and of template DNA into Xenopus oocytes, the mechanisms of transcription factor-mediated transcriptional activation of promoters within a chromatin environment can be examined in vivo (Almouzni and Wolffe, 1993; Wong et al., 1995). Presumably, nuclear factor(s) capable of binding to chromatin-reconstituted β-MyHC promoter and causing the disruption of its nucleosomal structure, will be the key regulator(s) responsible for the re-expression of this gene in the hypertrophic heart.

The mammalian SWI-SNF complex is heterogenous, with some subunits preferentially expressed in heart (Wang et al., 1996a and b). Four histone acetyltransferases (HAT) have been identified, and more HATs await discovery (Yang et al., 1996; Ogryzko et al., 1996). In one of them, P/CAF, its mRNA is most abundant in cardiac and skeletal muscle (Yang et al., 1996). Therefore, it is speculated that different tissues may adopt different chromatin remodelling mechanisms in gene regulation. An established cardiac gene chromatin remodelling model should be useful for future research.
Fig. 25 Chromatin remodelling of the cardiac β-MyHC gene. (also see Future Experiments for details) In fetal stage, the β-MyHC gene promoter is organized in three positioned nucleosomes. The second nucleosome (the transparent circle) is DNase I hypersensitive, allowing all transcription factors to bind to their binding sites. The DNase I hypersensitivity of the β-MyHC gene promoter is lost in adult stage, which is probably caused by the effect of TR. The formation of compacted nucleosome structure excludes the transcription factors from the β-MyHC gene promoter. The molecular mechanism of chromatin remodelling in normal or hypertrophic hearts could be investigated, by using the (1) in vitro and (2) in vivo chromatin assembly systems. Therefore, a cardiac-specific chromatin remodelling model could be established (3).


Ball, D. J., Gross, D. S., and Garrard, W. T. (1983) 5-Methylcytosine is localized in nucleosomes that contain histone H1. Proc Natl Acad Sci USA 80, 5490-5494


Bednarik, D. P. Duckett, C., Kim, S. U. Perez, V. I., Griffis, K.,
inhibits binding of NF-κappa B proteins to the HIV-1 long terminal
repeat cognate DNA motifs. New Biol 3, 969-976

methylation in CTF and SP1 recognition sites of an HSV tk promoter:
Effects on transcription in vivo and on factor binding in vitro.
Nucleic Acids Res 17, 10179-10190

activation of human nuclear genes in stable heterocaryons. Cell
32, 1171-1180

Similarities in DNA-binding preferences of MyoD and E2A protein
complexes revealed by binding site selection. Science 250, 1104-
1110

Boheler, K. R., Carrier, L., Chassagne, C., de la Bastie, D.,
Mercadier, J. J. (1991) Regulation of myosin heavy chain and actin
isogenes expression during cardiac growth. Mol. Cell. Biochem. 104,
101-7

Boheler, K. R., Chassagne, C., Martin, X., Wisnewsky, C., Schwartz,


arginine-rich histone kernel. Cell 8, 333–347


enhancer factor 1 disruption by a retroviral gene trap leads to heart defects and embryonic lethality in mice. *Genes Dev.* 8, 2293-2301


Georgakopoulos, T., and Thireso, G. (1992) Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription. EMBO J. 11, 4145-4152


Gossett, L. A., Kelvin, D. J., Sternberg, E. A., and Olson, E. N.


152


Kariya, K., Karns, L., and Simpson, P. C. (1994) An enhancer core element mediates stimulation of the rat \( \beta \)-myosin heavy chain promoter by an \( \alpha_1 \)-adrenergic agonist and activated \( \beta \)-protein kinase C in hypertrophy of cardiac myocytes. J. Biol. Chem. 269, 3775-3782

Kaushal, S., Schneider, J. W., Nadal-Ginard, B., and Mahdavi, V. (1994) Activation of the myogenic lineage by MEF2A, a factor that induces and cooperates with MyoD. Science 266, 1236-1240


Lee, J. W., Choi, H.-S., Gyuris, J., Brent, R., and Moore, D. D. (1995) Two classes of proteins dependent on either the presence or absence of thyroid hormone for interaction with thyroid hormone receptors. Mol. Endocrinol. 9, 243-254


Lee, W., E. deBoer, and F. Grosveld.: The human β-globin gene 3' enhancer contains multiple binding sites for an erythroid-specific protein.


is developmentally and hormonally regulated. J. Biol. Chem. 259, 6437-6446


Nakatsuji, Y., Hidaka, K., Tsujino, S., Yamamoto, Y., Mukai, T., Yanagihara, T., Kishimoto, T., and Sakoda, S. (1992) A single MEF-2 site is a major positive regulatory element required for transcription of the muscle-specific subunit of the human phosphoglycerate mutase gene is skeletal and cardiac muscle


Pazin, M. J., Sheridan, P. L., Cannon, K., Cao, Z., Keck, J. G.,


Peterson, C. L. and Herskowitz, I. (1992) Characterization of the yeast SWI1, SWI2, and SWI3 genes, which encode a global activator of transcription. *Cell* 68, 573-583


adaptation of cultured cardiac cells. *J. Biol. Chem.* **267**, 10551-10560


Trends Biochem Sci. 12, 108-112


of GAGA transcription factor. *Nature* 367, 525-532


Wang, W., Cote, J., Xue, Y., Zhou, S., Khavari, P. A., Biggar, S.


Wong, J., Shi, Y.-B. and Wolffe, A. P. (1995) A role for nucleosome assembly in both silencing and activation of the Xenopus TRBA gene by the thyroid hormone receptor. Genes Dev. 9, 2696-2711


APPENDIX
A1. Fluorescence in situ hybridization for chromosome localization of hGATA-4

Materials and Methods

This experiment was done by Henry Heng. I provided the probe and wrote the manuscript (Huang et al., 1996).

The probe used in this study was a 1.1kb DNA fragment containing the unique N-terminal portion of the human GATA-4 (Huang et al., 1995). The probe was biotinylated with dATP using the BRL BioNick labeling kit. Fluorescence in situ hybridization (FISH) was performed as previously described (Heng et al., 1992). Briefly, metaphase spreads were baked at 55°C for 1 hour. After RNase A treatment, the slides were denatured at 70°C, and dehydrated with ethanol. After hybridization with the denatured probe, slides were washed, followed by detection, as well as amplification of signals, as previously described (Heng et al., 1994). The FISH signals and the DAPI banding pattern were visualized in one single operation by switching microscope filters.

Results and Discussion

It would be interesting to determine the chromosome localization of the hGATA-4. This experiment would allow us to see whether there is possible linkage of the GATA-4 gene to the candidate regions in genome relating to congenital heart diseases. DAPI banding was used to identify the specific chromosome to which the probe hybridized as chromosome 8. One hundred mitotic figures were checked, and the hybridization frequency was 80%. A total of
10 mitotic figures were photographed, of which one example is shown in Fig. 26A and 26B. The locations of the probe-hybridization signals are schematically summarized in Fig. 26C. The data indicate that the human GATA-4 gene is located on the chromosome 8 region p22-23.1. However, some weak signals in the other chromosomes were also observed, indicating that these might be the chromosomal locations of the GATA-4-related genes such as GATA-5 and GATA-6 (Laverriere et al., 1994).
Fig. 26 FISH mapping of the human GATA-4 gene. (A) Arrow indicates location of FISH signal. (B) The same metaphase as (A) but stained with DAPI for chromosome identification. (C) Summary of FISH data for human GATA-4 on the chromosome 8. Each dot represents a double fluorescent signal on this chromosome.
A2. Chloramphenicol Acetyl Transferase Assay Counts

Numbers are expressed as % incorporation =

\[
(1,3-\text{Acetylated Chloramphenicol})(100)*
\]

Total Counts

*All values have been normalized for β-Galactosidase activity.

For Fig. 20

<table>
<thead>
<tr>
<th># CAT Constructs</th>
<th>Trans. #1</th>
<th>Trans. #2</th>
<th>Trans. #3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Myo</td>
<td>Fib</td>
<td>Myo</td>
</tr>
<tr>
<td>pSV40</td>
<td>73.08</td>
<td>66.37</td>
<td>80.44</td>
</tr>
<tr>
<td>β-294 CAT</td>
<td>24.12</td>
<td>0.93</td>
<td>33.46</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th># CAT constructs</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Myo</td>
<td>Fib</td>
</tr>
<tr>
<td>pSV 40</td>
<td>76.79</td>
<td>71.48</td>
</tr>
<tr>
<td>β-294 CAT</td>
<td>28.76</td>
<td>1.21</td>
</tr>
</tbody>
</table>

For Fig. 21

<table>
<thead>
<tr>
<th># CAT</th>
<th>Cardiomyocytes</th>
<th>C2C12 Myotubes</th>
<th>Cardiomyoblasts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trans. #1</td>
<td>Trans. #2</td>
<td>Trans. #3</td>
</tr>
<tr>
<td>TK</td>
<td>36.17</td>
<td>39.33</td>
<td>27.63</td>
</tr>
<tr>
<td>E+TK(+)</td>
<td>48.1</td>
<td>37.34</td>
<td>44.02</td>
</tr>
<tr>
<td>E+TK(-)</td>
<td>29.38</td>
<td>28.88</td>
<td>31.24</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th># TK</th>
<th>Cardiomyocytes</th>
<th>C2C12 Myotubes</th>
<th>Cardiomyoblasts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Trans. #1</td>
<td>Trans. #2</td>
<td>Trans. #3</td>
</tr>
<tr>
<td>E(+)/TK</td>
<td>127</td>
<td>95</td>
<td>158</td>
</tr>
<tr>
<td>E(-)/TK</td>
<td>81</td>
<td>73</td>
<td>112</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>% TK</th>
<th>C2C12</th>
<th>Myo</th>
<th>Fib</th>
<th>C2C12</th>
<th>Myo</th>
<th>Fib</th>
</tr>
</thead>
<tbody>
<tr>
<td>E(+)/TK</td>
<td>312</td>
<td>126.67</td>
<td>92</td>
<td>24.51</td>
<td>31.5</td>
<td>19.97</td>
</tr>
<tr>
<td>E(-)/TK</td>
<td>136.67</td>
<td>88.7</td>
<td>112.67</td>
<td>27.4</td>
<td>20.6</td>
<td>14.84</td>
</tr>
</tbody>
</table>
For **Fig. 22A**

<table>
<thead>
<tr>
<th>CAT constructs</th>
<th>Trans. #1</th>
<th>Trans. #2</th>
<th>Trans. #3</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>β-294 CAT</td>
<td>38.52</td>
<td>30.27</td>
<td>36.04</td>
<td>34.94</td>
<td>4.23</td>
</tr>
<tr>
<td>mβ-294 CAT</td>
<td>11.81</td>
<td>10.89</td>
<td>13.08</td>
<td>11.86</td>
<td>1.26</td>
</tr>
</tbody>
</table>

For **Fig. 22C**

<table>
<thead>
<tr>
<th>CAT constructs</th>
<th>Trans. #1</th>
<th>Trans. #2</th>
<th>Trans. #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>GATA-4 + β-294</td>
<td>8.89</td>
<td>11.03</td>
<td>6.26</td>
</tr>
<tr>
<td>pMT-2 + β-294</td>
<td>1.47</td>
<td>2.02</td>
<td>1.13</td>
</tr>
<tr>
<td>GATA-4 + mβ-294</td>
<td>1.1</td>
<td>1.58</td>
<td>0.9</td>
</tr>
<tr>
<td>pMT-2 + mβ-294</td>
<td>1.32</td>
<td>1.6</td>
<td>0.99</td>
</tr>
<tr>
<td>β-294 CAT</td>
<td>1.35</td>
<td>1.87</td>
<td>0.84</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>% β-294 CAT</th>
<th>Trans. #1</th>
<th>Trans. #2</th>
<th>Trans. #3</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>GATA-4 + β-294</td>
<td>659</td>
<td>590</td>
<td>745</td>
<td>665</td>
<td>77.65</td>
</tr>
<tr>
<td>pMT-2 + β-294</td>
<td>109</td>
<td>108</td>
<td>134</td>
<td>117</td>
<td>14.73</td>
</tr>
<tr>
<td>GATA-4 + mβ-294</td>
<td>81</td>
<td>84</td>
<td>107</td>
<td>90.7</td>
<td>14.22</td>
</tr>
<tr>
<td>pMT-2 + mβ-294</td>
<td>97.7</td>
<td>85.5</td>
<td>118</td>
<td>100.5</td>
<td>16.41</td>
</tr>
</tbody>
</table>
**A3. Northern blot analysis of the hamster cardiac MyHC genes** (This work was performed by Dr. R. X. Wang (Wang, 1994).

(A) **Fig. 27 Northern blot analysis of the hamster cardiac α-** (A) and β-(B) MyHC gene** In this analysis, part of the 3' untranslated region of each gene was used as a probe. The length of the probe is 398 bp and 411 bp for the α- and β-MyHC gene, respectively. (A) The expression of the α-MyHC gene was low in 2 day heart (lane 1), and high in 10 day (lane 2) and 20 day (lane 3) ventricles. The highest expression was observed in 180 day atrial sample (lane 4). (B) The expression of the β-MyHC gene was high at 2 day old ventricle (lane 1) and declined significantly by 10 days of age (lane 2); low level of expression was observed in 180 day atrial sample (lane 3).