Maastricht University

Cardiac specific gene expression of the regulatory myosin light chains

Citation for published version (APA):

Doevendans, P. A. F. M. (1997). *Cardiac specific gene expression of the regulatory myosin light chains*. [Doctoral Thesis, Maastricht University]. Universiteit Maastricht. https://doi.org/10.26481/dis.19971212pd

Document status and date:
Published: 01/01/1997

DOI:
10.26481/dis.19971212pd

Document Version:
Publisher's PDF, also known as Version of record

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 23 Apr. 2024

Cardiac specific gene expression of the regulatory myosin light chains

Legend to the picture:
Transiently transfected neonatal rat cardiomyocytes assemble normal sarcomeres. Primary cultures of neonatal rat cardiomyocytes were transfected with expression plasmids harboring human beta myosin heavy chain cDNA. Two days after transfection the cells were fixed and immunstained using an epitope specific antibody recognizing the human protein. The fluorochrome conjugated to the secondary antibody was visualized via epifluorescence and confocal microscopy.

© 1997 P.A.F.M. Doevendans, Maastricht, The Netherlands
All rights are reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording or any information storage or retrieval system, without permission in writing from the copyright owner.

ISBN 90 5278 235 0

Cardiac specific gene expression of the regulatory myosin light chains

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit Maastricht,
op gezag van de Rector Magnificus, Prof. mr. M. J. Cohen,
volgens het besluit van het College van Decanen
in het openbaar te verdedigen
op vrijdag 12 december 1997 om 16.00 uur

door

Pieter A.F.M. Doevendans

★
UPM
UNIVERSITAIRE PERS MAASTRICHT

Promotor
Prof. dr. H.J.J. Wellens

Co-promotor
Dr. M. van Bilsen

Promotiecommissie
Prof. dr. H. Struijker Boudier (voorzitter)
Prof. dr. M. Daemen
Prof. dr. W.H. van Gilst (Rijksuniversiteit Groningen)
Prof. dr. A. van der Laarse (Rijksuniversiteit Leiden)
Prof. dr. F. Ramaekers

De uitgave van de proefschrift werd mede mogelijk gemaakt door financiële steun van de Nederlandse Hartstichting en de Stichting RESCAR.

Financiële steun werd verder geleverd: Asta Medica BV, Bayer BV, Byk Nederland BV, Lorex Synthélabo BV, Menarini Benelux BV, Merck Sharp & Dohme BV, Sanofi Winthrop BV, Novartis BV, Bristol Myers Squibb BV, Rhône-Poulenc Rorer BV, Park Davis BV, Astra BV.

Voor mijn ouders

Table of contents

Abbreviations 9

CHAPTER 1
General introduction 11

CHAPTER 2
Transcription factors and the cardiac gene programme. 21

CHAPTER 3
Hunting down nucleic acid binding proteins in the cardiovascular system. 53

CHAPTER 4
Developmental and tissue restricted expression of a murine spliceosome associated protein gene mSAP49 during cardiogenesis. 85

CHAPTER 5
Characterization of embryonic stem cell derived cardiomyocytes: a comparative study of *in vivo and in vitro* developed murine cardiomyocytes. 99

CHAPTER 6
Cloning of the murine atrial myosin light chain-2 gene: a member of an evolutionary conserved family of contractile proteins. 119

CHAPTER 7
The *in vitro* characterization of the murine atrial myosin light chain 2 promoter. 133

CHAPTER 8
The regulatory atrial myosin light chain gene is re-expressed in ventricular myocytes under hypertrophic conditions. 159

CHAPTER 9
Inhibition of the atrial myosin light chain-2 promoter in cultured cardiomyocytes by nuclear receptors. 181

CHAPTER 10
General discussion 199

Summary 207
Samenvatting 209
Acknowledgements 211
Publications 213

Abbreviations

ANF	Atrial natriuretic factor
atRA	All trans retinoic acid
bp	Basepair
BSA	Bovine serum albumine
Cdk	Cyclin dependent kinase
cDNA	Copy DNA
CMV	Cytomegalovirus
CRBP II	Cellular retinol binding protein type II
CRE	cAMP responsive element
CT-1	Cardiotrophin-1
DEPC	Diethyl Pyrocarbonate
DMSO	Dimethyl sulphoxide
dpc	Days post coitum
DR1	Direct repeat (spaced by 1 nucleotide)
EB	Embroid body
EBDC	Embryoid body derived cardiomyocytes
ERE	Estrogen responsive element
ES	Embryonic stem (cell)
ET-1	Endothelin1
FITC	Fluorescein isothiocyanate
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GFP	Green fluorescent protein
GMSA	Gel mobility shift assay
HLH	Helix-loop-helix
HTH	Helix-turn-helix
Id	Inhibitor of differentiation
Kb	Kilobases (1000 bps)
LCR	Locus control region
LEF1	Leucocyte enhancer-binding factor
LIF	Leukemia inhibitory factor
MADS	MCM 1, agamous, deficiens, SRF (family of proteins)
MAPKK	Mitogen activated protein kinase kinase
MCK	Muscle creatine kinase
MeCP1	Methyl CpG binding proteins
MEF	Myocyte-specific enhancer binding factor
MHC	Myosin heavy chain

Mkp-1	Mitogen-activated protein kinase phosphatase-1
MLC-1	Essential or alkali myosin light chain
MLC-2	Regulatory myosin light chain
mRNA	Messenger RNA
NABP	Nucleic acid binding proteins
PBS	Phosphate buffered saline
PBSAN	PBS with 1% BSA, and 0.2% non-idet P-40
PCR	Polymerase chain reaction
PE	Phenylephrine
POU	Pit 1, Oct 1, Unc. (family of proteins)
PPAR	Peroxisomal proliferator-activated receptors
RA	Retinoic acid
RACE	Rapid amplification of cDNA ends
RAR	Retinoic acid receptor
RSRF	Related to Serum Response Factor
RSV	Rous-Sarcoma virus
rtPCR	Reverse transcriptase polymerase chain reaction
RXR	Retinoid X receptor
SRE	Serum response element
SRF	Serum response factor
SSAP	Stage Specific Activator Protein
TAF	TATA binding protein associated factor
TBP	TATA binding protein
TnT	Troponin T
TR	Thyroid responsive element
TTNPB	Synthetic retinoic acid agonist
USF	Upstream Stimulating Factor
VDR	Vitamine D responsive element

Chapter

General introduction

The contraction of cardiac muscle cells requires the presence of functional units (sarcomeres), composed of a specialized set of contractile proteins. In vertebrates, the ultrastructure and protein composition of these sarcomeres is remarkable similar among different striated muscle types. However underlying this apparent morphological uniformity there is a high degree of molecular diversity, due to the existence of multiple isoforms of each myofibrillar component e.g. α and β myosin heavy chain, atrial and ventricular myosin light chain 2.[1] This thesis focuses on the expression of the cardiac regulatory myosin light chains. Therefore, a brief introduction is given on the role of the ventricular and atrial myosin light chain 2 isoforms on sarcomeric structure and function.

Sarcomere structure

The major components of the sarcomere are myosin, actin, tropomyosin and the three components of troponin (T,I,C, figure 1). These contractile proteins have specific roles in the interactions that enable muscle to relax and contract. The nomenclature used for the troponin proteins is derived from their function. Troponin I inhibits (I) the interactions between actin and myosin, troponin T binds the troponin complex to tropomyosin (T) and troponin C contains the calcium (C) binding sites that regulate contraction. Actin, troponin and tropomyosin constitute the thin filament, whereas myosin is the major protein of the thick filament. Every myosin molecule is a hexamer containing 2 myosin heavy chains and two pairs of light chains. The myosin molecule can be morphologically distinguished in a tail, neck and head region. The model of contraction based on current molecular knowledge, schematically represented in figure 1, involves movement of the myosin neck at the hinge region where neck and tail are connected. The change of angle between tail and neck will move the head of the myosin molecule along the thin filaments.[2] The head of the myosin molecule can hydrolyze ATP and bind to actin, forming cross bridges between the thin and thick filaments. The light chains are positioned in tandem at the neck of the myosin heavy chain molecule. The regulatory light chain is attached to the hinge region and therefore can play an important role in changes in the myosin neck during contraction. The essential light chain interacts with the head of the myosin molecule.[3]

Myosin light chain diversity and nomenclature

The light chains can be classified as the calcium binding regulatory light chain (MLC-2) and the essential light chain (MLC-1, MLC-3). Terminology can be confusing as light chain classification used to be based on the chemical properties of the proteins (table 1). The light chains 1 and 3 can be removed from the myosin molecules by exposure to a high pH. Therefore these light chains have

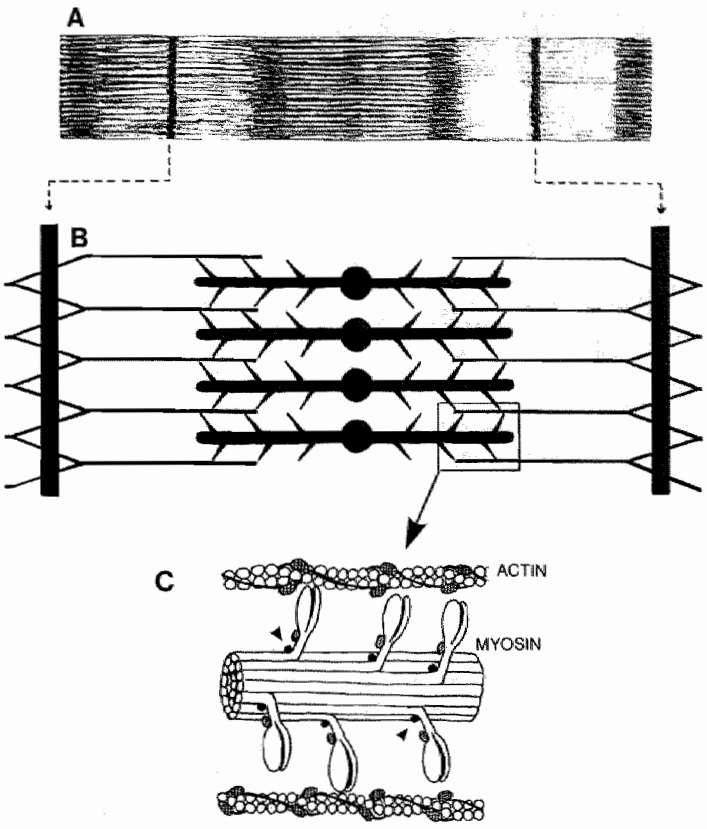


Figure 1. Schematic representation of striated muscle. A) The microscopic aspect of the cross striations and the various structural units composing the sarcomere in a relaxed state. B) Schematic drawing showing the thin and thick filament and their position at the same magnification as panel A. C) Proteins assembled in the sarcomere with the myosin crossbridges interacting with actin. The position of the MLC-2 is indicated at the neck of the myosin molecules (arrow). Note the position of the regulatory myosin light chain at the hinge of the tail and neck of the myosin heavy chain.

been classified as alkali myosin light chains and abbreviated as aMLC proteins. The regulatory light chain proteins can be separated from myosin by non-denaturing agents as 5,5'-dithiobis-(2-nitrobenzoic acid) and ethylenediamine tetra-acetate. Different MLC-2 isoforms have been identified in atrial, ventricular, fast skeletal (-2f) and smooth muscle and non muscle cells (-2sm). These isoforms represent a multigene family. In contrast, the essential light chain 1 and 3 are derived from one gene by developmentally regulated alternative splicing.

Table 1. Mammalian myosin light chains

Gene	Nomenclature	Expression
Essential, alkali		
MLC 1/3 fast	MLC-1f, MLC-3f	Fast skeletal muscle (isoforms 1 and 3 through alternative splicing (19)
MLC 1 slow-a	MLC-1sa	Slow skeletal muscle, smooth muscle and non muscle (20)
MLC 1 slow/ventricular	MLC-1sb, MLC-1s, MLC-1v	Slow skeletal muscle, ventricle (19)
MLC 1 embryonic/atrial	C-1, MLC-1a	Developing skeletal muscle/atrium (19)
Regulatory, phosphorylatable		
MLC-2 fast	MLC-2f	Fast skeletal muscle (21)
MLC-2 slow/ventricular	MLC-2s, MLC-2v	Slow skeletal muscle ventricle (13)
MLC-2 mandibular	MLC-2m, MYL5	Mandibular muscle, fetal skeletal muscle (22)
MLC-2 atrial	MLC-2a	Atrium (18)
MLC-2 sm	MLC- C20	Smooth muscle and non-muscle (3)

Throughout this thesis the atrial and ventricular tissue specific isoforms of the regulatory MLC-2 protein are referred to as MLC-2a and MLC-2v, respectively.

Myosin light chain function

In smooth muscle cells the function of the MLC-2 protein is well defined. Smooth muscle cell contraction is initiated when calcium binds to calmodulin. The calcium-calmodulin complex activates the MLC kinase, which phosphorylates the smooth muscle MLC-2. These reactions are mandatory for initiation of contraction. Whether the MLC-2 plays such a dominant role in cardiac myocyte contraction is unclear. For optimal cardiomyocyte contraction however both calcium handling and MLC-2 phosphorylation are important. The troponin complex and tropomyosin inhibit actin and myosin interaction at low calcium levels. Increased calcium levels will reverse this inhibitory effect. Calcium binding of the troponin C protein reduces the Troponin I affinity for actin and allows binding of myosin to actin. Calcium is not required for the interaction between actin and myosin *per se*. The light chains (essential and regulatory) are capable of Ca^{2+} binding through their structural domains. The light chains contain 4 so called E-F hands motifs (like troponin C), composed of two α helical regions and a connecting loop. The conformation of amino acids is such that Ca^{2+} can be

bound in the center of the domain. Binding of Ca^{2+} to MLC-2 protein results in a conformational change of the protein and improved contractile function. The proteins have a high level of homology with calmodulin, a ubiquitously expressed protein involved in kinase activation after calcium binding. One of the Ca^{2+}/calmodulin-activated kinases is the MLC kinase. Three pathways of calcium regulation of contraction emerge, namely through interaction with troponin I, direct calcium binding of the MLC, and phosphorylation of the MLC-2 through the MLC-kinase.

The functional role of the light chains in the sarcomere has gradually been resolved. *In vitro* studies have shown that removal of both light chains (essential and regulatory) reduces contraction velocity, without a reduction in energy consumption. Therefore the light chains appear to play a role in efficient conversion of energy in actin-myosin movement.[4] The functional consequences of atrial MLC-1 re-expression on myosin cycling kinetics have been studied in human demembranated (skinned) fibers from the right ventricular infundibulum of patients with congenital cardiac disease. *Ex vivo* maximal shorting velocity and isometric tension production was measured.[5] A near linear positive relation was found between the level of MLC-1a expression in ventricular tissue on one hand and shortening velocity and isometric tension production on the other. Thus re-expression of MLC-1a improves the function of ventricular skinned fibers. In these studies the potential re-expression of the atrial regulatory light chain (MLC-2a) was not determined. However, the introduction of molecular techniques has provided interesting information on regulatory light chain function, especially with respect to smooth muscle myosin. Studies in cardiomyocytes are complex because of the terminal differentiation of this cell type. Therefore most information is obtained in alternative muscle cell types. Mutation studies of the MLC-2sm protein, showed that phosphorylation of the protein at the serine residue at position 19 is crucial for function. The mechanism of function enhancement is phosphate group dependent (independent of amino-acid position) and is not explained by charge neutralization (replacement of positively charged amino acids).[6] Chamber specific MLC-2 protein replacement through transgenic technology is a new tool to study protein function *in vivo*. This strategy was followed to replace the MLC-2 protein in the atrial chambers by introducing the MLC-2v cDNA driven by the α MHC promoter. The MLC-2a protein was completely substituted by MLC-2v and loss of cardiac function was measured *ex vivo*.[7,8] These data are in line with the functional improvement seen in the presence of MLC-1a in human ventricular myocytes.

Pathology

In humans, cardiac pathology is often accompanied by altered gene expression. This has been shown for patients with congenital cardiac disease like tetralogy

of Fallot and double outlet left ventricle. In ventricular tissue of these patients re-expression of the essential atrial MLC (MLC-1a) is found, whereas in normal human heart expression of this isoform is confined to the atrial chambers only.[5]

Molecular analysis has identified mutations in various genes resulting in hypertrophic cardiomyopathy. Recently mutations in the ventricular essential and regulatory MLC were found in families with hypertrophic cardiomyopathy.[9] Missense mutations in the MLC gene are the cause in less then 1% of the cardiomyopathy families. In addition a skeletal myopathy was diagnosed in skeletal muscle biopsies of these patients. An interesting clinical feature in this hypertrophic cardiomyopathy form is the midventricular gradient which is present in all affected individuals. Apparently the mutation defines the predilection site where hypertrophy is most pronounced. In the presence of the MLC mutation this leads to papillary muscle hypertrophy and midcavitary obstruction. This finding supports the theory of geno-phenotype relation as the pattern of hypertrophy and obstruction was comparable in unrelated families with different mutations. The mutations found in the MLC-2v gene were all located in close proximity to the phosphorylation site.

The murine regulatory myosin light chain-2 genes

The MLC-2v gene has been studied intensively for several years by Chien and co-workers.[12] This light chain proved to be an excellent tool for many different molecular approaches to study cardiac transcription and protein function. The expression pattern of the gene appeared to delineate ventricular chamber formation, both at the RNA and protein level. The regulatory sequence of this gene, is responsible for tissue specific expression as transcription is limited to ventricular myocytes and slow twitch skeletal muscle.[13] The endogenous promoter is stimulated under hypertrophic conditions.[14] Furthermore the proximal promoter region has been used successfully to identify and clone tissue specific transcription factors.[15] In transgenic experiments the proximal promoter was sufficient to drive transgene expression to the ventricular compartment. This has been established in studies where the MLC-2v promoter was used to drive the expression of luciferase, galactosidase[16] and the positive dominant mutation of the human RAS gene.[17] Also, the atrial isoform proved to be an excellent molecular marker albeit that initial expression of the gene is throughout the entire heart tube, including the ventricular pre-chamber. Later during murine development (day 12 post coitum) expression becomes restricted to the atrial tissue. Therefore the promoter of the MLC-2a gene is anticipated to direct atrial specific gene expression. Accordingly, in theory the atrial tissue specific promoter could be used in the future to direct gene expression to the atrial compartment.

Aims of the Study

Based on the expression pattern of the MLC-2a gene, the MLC-2a regulatory sequence should be applicable to induce atrial myocyte specific gene expression. To test this hypothesis the MLC-2a promoter was cloned and characterized. One of the purposes of this study was to identify DNA elements and nuclear transcription factors that are responsible for the up- and downregulation of MLC-2a expression in cardiac myocytes during development and under, physiologic and pathologic conditions (figure 2). We specifically searched for regulatory DNA within the MLC-2a gene required for expression in atrial cardiomyocytes. The experiments on terminally differentiated neonatal atrial cardiomyocytes would require the sacrifice of large numbers of rats for cell isolation, therefore alternative sources of atrial cardiomyocytes were validated. In addition the behavior of the endogenous MLC-2a gene in ventricular tissue in human cardiac disease was analyzed.

Outline of the thesis

Chapter 2 provides a closer look at transcription relutation of various genes coding for cardiac contractile proteins. In that chapter the different DNA elements capable of protein binding and the role of various transcription factors in cardiac development and cardiac specific gene expression are discussed. In chapter 3 the methods to identify new DNA (transcription regulation) and RNA (splicing) binding proteins are presented together with a discussion of their applications in the cardiovascular field to illustrate the value of these methods. Many of these techniques were applied in subsequent chapters to study the regulation of the murine ventricular and atrial MLC-2 genes.
In the process of transcription regulation studies of the MLC-2v promoter we cloned a splice factor, mSAP49. This splice factor itself has an interesting pattern of expression during development. The timing and pattern of mSAP 49 expression is suggestive for a role in isoform switches in differentiating cardiomyocytes (chapter 4).

The MLC-2v and MLC-2a expression patterns provide molecular markers at the RNA and protein level to follow the process of chamber specification during cardiogenesis. This is important not only to study developmental cardiac pathology and congenital abnormalities, but also extremely useful in the validation of the differentiation state of *in vitro* developed cardiomyocytes. A potential alternative source for cardiomyocytes is provided by the embryonic stem cell derived embryoid bodies. The development of embryoid bodies could be a valid in vitro system for early cardiogenesis. The validation of embryoid body derived cardiomyocytes is presented in chapter 5. Here we compared the expression of the

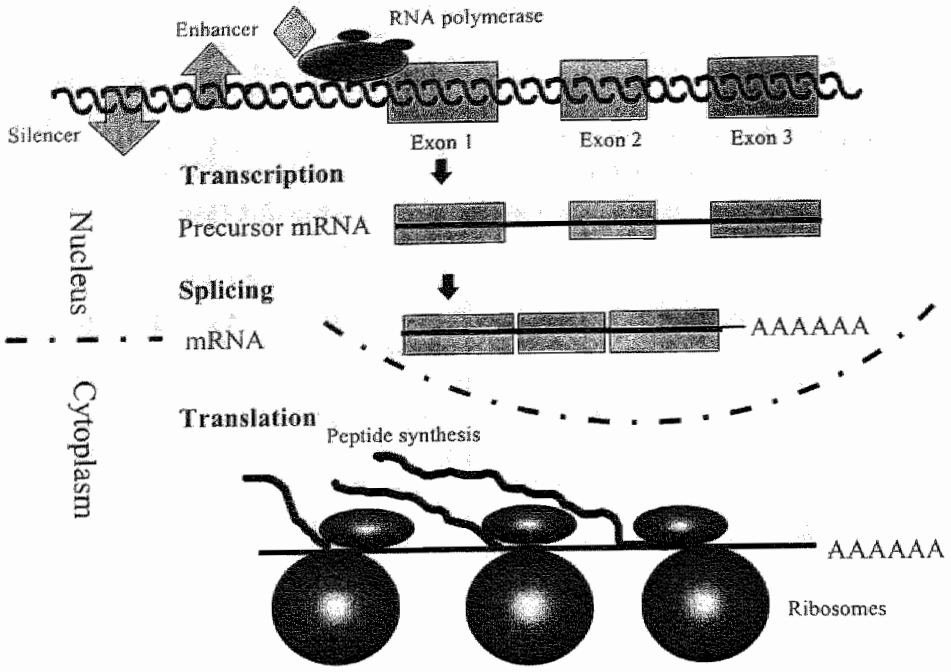


Figure 2. Schematic diagram showing the general gene structure. The gene can be divided in a regulatory sequence including transcription enhancing and silencing elements. The transcription start site is indicated by the position of the transcription initiation complex including the enzyme RNA polymerase. The remainder of the gene, provides the template for RNA formation during the process of transcription, and starts at exon 1. The precursor RNA molecule is converted into messenger RNA (mRNA) through splicing. The final mRNA contains only the exons which form the coding sequence. The mRNA leaves the nucleus and the nucleic acid code is translated at the site of the ribosomes into amino-acids. The linkage of amino-acids results in protein peptide formation.

MLC-2 genes at the protein level, together with a functional analysis of ion channel proteins. The immunostaining of cells following the whole cell patch clamp procedure is a new technique to study individual cardiomyocytes.

The MLC-2v gene has been used to identify important cardiac transcription factors. Furthermore the promoter of this gene has been used successfully to mediate cardiac specific gene expression both in transfection and transgenic experiments. Stimulated by the results obtained in studying the MLC-2v gene, the atrial isoform gene was searched for and identified. The cloning of the near complete cDNA of the gene was previously reported.[18] In chapter 6 the structure of the entire coding region is presented including approximately 2000 basepairs (bp) of the 5' and 700 bp of the 3' flanking region. The structure of the gene is compared to other MLC genes from various species.

In chapter 7 the regulatory sequence of the murine atrial myosin light chain-2 gene (5' flanking region) was characterized by *in vitro* studies. To locate DNA regulatory sequences that confer atrial myocyte specific gene expression, promoter activity was assessed in different cell types. For these studies the deletion fragments of the MLC-2a promoter were cloned into reporter gene (luciferase) vectors and transfected into rat neonatal atrial and ventricular cardiomyocytes, manipulated atrial myocytes (AT1), and skeletal myocytes (C2C12) as well as monkey kidney cells (COS). The COS cells were used as a representative of the non muscle cell types.

During transfection studies we found re-expression of the endogenous MLC-2a gene in cultured rat neonatal ventricular myocytes (chapter 8). The expression of the MLC-2a regulatory sequence was investigated under various serum conditions. In addition various transfection experiments were performed in the presence of hypertrophy inducing factors and estrogen. Finally human tissue samples were included for RNA and protein analysis.

Within the regulatory MLC-2a sequence several DNA elements were identified that potentially could interact with activated retinoic acid receptors. The analysis of one selected retinoic acid responsive elements through gel mobility shift assays, mutation analysis and transient transfection assays is presented in chapter 9.

In the last chapter (chapter 10) implications of these studies of regulatory sequences, nucleic acid binding proteins, and the contractile proteins of the murine regulatory MLC are discussed, addressing directions for future research and potential applications.

Literature

1. Schiaffino S, Reggiani C: Molecular diversity of myofibrillar proteins: gene regulation and functional significance. *Physiol Rev* 1996;76:371-423
2. Katz AM: Contractile proteins, in Katz AM (ed): *Physiology of the heart*. ed-2nd. New York, Raven Press, 1992, 151-177
3. Trybus KM: Role of myosin light chains. *J Musc Res Cell Mot* 1994;15:587-594
4. Lowey S, Waller GS, Trybus KM: Skeletal muscle myosin light chains are essential for physiological speeds of shortening. *Nature* 1993;365:454-456
5. Morano M, Zacharzowski U, Maier M, Lange PE, Alexi Meskishvili V, Haase H, Morano I: Regulation of human heart contractility by essential myosin light chain isoforms. *J Clin Invest* 1996;98:467-473
6. Sweeney HL, Yang Z, Zhi G, Stull JT, Trybus KM: Charge replacement near the phosphorylatable serine of the myosin regulatory light chain mimics aspects of phosphorylation. *Proc Natl Acad Sci U S A* 1994;91:1490-1494
7. Palermo J, Gulick J, Ng W, Grupp IL, Grupp G, Robbins J: Remodeling the mammalian heart using transgenesis. *Cell Mol Biol Res* 1995;41:501-509

8. Gulick J, Hewett TE, Klevitsky R, Buck SH, Moss RL, Robbins J: Transgenic remodeling of the regulatory myosin light chains in the mammalian heart. *Circ Res* 1997;80:655-664
9. Spirito P, Seidman CE, McKenna WJ, Maron BJ: The management of hypertrophic cardiomyopathy. *N Engl J Med* 1997;336:775-785
10. Barton PJR: Molecular biology and the heart, in Yacoub M, Pepper J (eds): *1990-91 annual of Cardiac Surgery*. London, Current Science, 1991, pp 3-12
11. Doevendans PA, van Gilst WH, Laarse A, van Bilsen M: Molecular cardiology I: Gene transcription in the cardiovascular system. *Cardiologie* 1997;4:221-229
12. Chien KR: Genes and physiology: molecular physiology in genetically engineered animals. *J Clin Invest* 1996;98:S19-26
13. O'Brien TX, Lee KJ, Chien KR: Positional specification of ventricular myosin light chain 2 expression in the primitive murine heart tube. *Proc Natl Acad Sci U S A* 1993;90:5157-5161
14. Shubeita HE, Martinson EA, van Bilsen M, Chien KR, Brown JH: Transcriptional activation of the cardiac myosin light chain 2 and atrial natriuretic factor genes by protein kinase C in neonatal rat ventricular myocytes. *Proc Natl Acad Sci U S A* 1992;89:1305-1309
15. Zhu H, Nguyen VT, Brown AB, Pourhosseini A, Garcia AV, van Bilsen M, Chien KR: A novel, tissue-restricted zinc finger protein (HF-1b) binds to the cardiac regulatory element (HF-1b/MEF-2) in the rat myosin light-chain 2 gene. *Mol Cell Biol* 1993;13:4432-4444
16. Ross RS, Navankasattusas S, Harvey RP, Chien KR: An HF-1a/HF-1b/MEF-2 combinatorial element confers cardiac ventricular specificity and established an anterior-posterior gradient of expression. *Development* 1996;122:1799-1809
17. Hunter JJ, Tanaka N, Rockman HA, Ross J, Jr., Chien KR: Ventricular expression of a MLC-2v-ras fusion gene induces cardiac hypertrophy and selective diastolic dysfunction in transgenic mice. *J Biol Chem* 1995;270:23173-23178
18. Kubalak SW, Miller-Hance W, O'Brien T, Dyson E, Chien K: Chamber-specification of atrial myosin light chain-2 expression precedes septation during mouse cardiogenesis. *J Biol Chem* 1994;269:16961-16970.
19. Barton PJ, Buckingham ME: The myosin alkali light chain proteins and their genes. *Biochem J* 1985;231:249-261.
20. Hailstones DL, Gunning PW: Characterization of human myosin light chains Isa and 3nm: implications for isoform evolution and function. *Mol Cell Biol* 1990;10:1095-1104.
21. Nudel U, Calvo JM, Shani M, Levy Z: The nucleotide sequence of a rat myosin light chain 2 gene. *Nucleic Acids Res.* 1984;12:7175-7186.

Chapter

Transcription factors and the cardiac gene programme

Pieter A. Doevendans MD, Marc van Bilsen PhD.
Published: in the International Journal of Biochemical Cell Biology 1996;28:387-403)

Abstract

During the past decade, major advances have been made in uncovering the mechanisms that switch genes on and off. Gene methylation and histones play an important role in gene (in)activation. Following gene activation, the initiation of transcription by RNA polymerase requires the assembly of multiple protein complexes on the promoter region of a gene. How a cell type-specific gene expression pattern can be induced is a key question in cardiovascular biology today. Member of the helix-loop-helix-family of the transcription factors play a dominant role in skeletal muscle formation. In cardiac muscle the situation is less obvious. Recent studies identified muscle transcription factors like MEF-2, TEF-1 and MNF, which are common to both the skeletal and cardiac muscle lineages. A few transcription factors, among which Nkx 2.5 and GATA-4, are expressed predominantly in the heart. The absence of master regulators in the heart points to the importance of interaction between ubiquitous factors and tissue restricted factors to initiate the cardiac gene programme and to lock these cells in their differentiated state. The recent development of murine transgenic and gene-targeting technology provides tools to study the role of mammalian transcription factors *in vivo*. Interesting cardiac phenotypes are found in gene targeted mice, indicating a crucial role for retinoic acid and homeobox genes in murine cardiogenesis.

Introduction

The heart is a fascinating and complex organ, where muscle activity is controlled by an extensive specialized conduction system. This sophisticated interaction provides a contraction pattern which allows an optimal pump efficiency. The repetitive contractile function of the heart requires a set of unique proteins involved in ion transport, contraction, and metabolism. The fully grown heart is capable of adapting to changing physical conditions, like elevated blood pressure, to maintain blood flow to remote organs. On a short term basis, functional adaptations are accomplished by changes in the neural stimulation of the heart and by humoral factors. More sustained changes are associated with alterations in gene expression and, hence, require adjustment of the transcriptional level of specific genes. A general feature of this 'remodelling' process is the re-expression of a panel of foetal genes (review [1]).

Different mechanisms contribute to gene activation and transcription in a tissue restricted manner. The skeletal muscle phenotype appears to be determined at large by myogenic transcription factors capable of induction of the muscular gene programme in non-muscle cells (review [2]). In skeletal myocytes, differentiation and cell division are mutually exclusive.[2] The cardiac gene programme, however, is different in many ways. Cardiomyocytes differentiate early in embryogenesis and, despite the presence of a functional contractile apparatus, they are still capable of DNA replication and cell division. Following the attenuation of cell division shortly after birth, cardiomyocytes are considered to be terminally differentiated and hypertrophy is the only adaptive response available for cardiomyocytes.[3] Although it has been shown that several transcription factors are crucial for normal cardiogenesis, no factors have been found that can initiate the cardiac specific gene programme in cell types other than cardiomyocytes. The formation of a four chamber organ from apparently simple blood isles early in embryonic development is a poorly understood phenomenon and a continuously challenging issue for developmental biologists. Gradually, the crucial mechanisms and key-players in cardiac determination and differentiation are being identified and through new molecular techniques, active manipulation of this process in different species is within reach (review [4]).

So far studies of the cardiac gene programme are hampered by the absence of stable cell lines that can be used for *in vitro* studies and transfection assays. The use of embryonic stem cell derived embryoid bodies to study early cardiogenesis *in vitro*, proved to be a valid system and provides an alternative source of cardiac precursor cells. The introduction of transgenes and gene targeting in oocytes and embryonic stem cells allows functional studies of different factors *in vivo* in transgenic organisms.[5,6]

In this review, we will follow the cardiac gene programme during cardiogenesis and cardiomyocyte differentiation. The process of gene activation and tran-

Table 1. Regulatory mechanism involved in tissue specific expression:

A) Tissue specific gene activation (8)
B) Tissue specific TBP associated factors? (150)
C) Tissue specific transcription factors (39)
D) Interaction between tissue specific and ubiquitous factors (41, 94)
E) Post-transcriptional regulation (31)
F) Post-translational modification (45)
G) Gradient in ligand or transcription factor expression (91)

Abbreviations: TBP: TATA binding protein.

scription in general will be discussed briefly. The muscle specific transcription factors and their characteristics, especially those that are relevant to the cardiac muscle lineage, are addressed in detail.

Regulation mechanisms of gene activation and initiation of transcription

Before *trans*-activating factors can contribute to transcription regulation, they have to interact with *cis*-regulatory DNA sequences of a target gene. Most genes, however, are packed in a highly organized structure, with extensive DNA-protein interaction. Therefore, the process of gene activation in which *cis*-elements are exposed to *trans*-activating factors has to precede transcription.

The organization of the supercoiled DNA in interphase nuclei is dependent on DNA binding and scaffolding proteins, of which the histones are the best known. The histones H2A, H2B, H3 and H4 are highly conserved between different species. These histones form an octameric core particle around which the DNA is wound twice. These nucleosomes are connected by linking DNA and additional linking histones, like H1. The linking histones are less conserved and can be tissue or stage specific.[7] For example, in avian nucleated red blood cells the H1 linking histone protein present in most cells is replaced by the histone H5 (A, table 1,[8]). The histones may stimulate transcription by presenting the DNA to *trans*-activating factors. Alternatively, they inhibit transcription by their interaction with regulatory elements. For instance, the promoter of the mouse mammary tumor virus gene is organized as an array of five positioned nucleosomes. In this configuration the promoter is inactive. Binding of the glucocorticoid receptor leads to nucleosome disruption and to transcription of the gene.[9] Clearly, the nucleosome structure contributes to transcriptional regulation by organizing the DNA in a reproducible, specific way (review [10-12]).
Methylation of CpG islands in promoter regions is a mechanism to modify gene activation during development and in a tissue restricted manner (A, table 1).

Methylation is a reversible process by which the binding of transcription factors is effectively blocked.[13] In addition, two proteins have been identified capable of binding methylated regions (MeCP1 and 2).[14] MeCP1 was found to repress transcription of genes with methylated promoter regions. The importance of methylation during development is further substantiated by the notion that homozygous knock-out (-/-) murine embryos lacking the enzyme DNA-methyltransferase die at mid-gestation 10 days post-coitum (dpc).[15,16]

Once a gene has been activated, it can be transcribed into RNA (figure 1, chapter 1). Transcription of genes in mammalian cells requires the binding of a complex of proteins to the TATA box, located just upstream from the transcription start site. Some genes lack the canonical TATA sequence (TATA(A/T)A(A/T) and have a modified TATA-like DNA element.[17] In several house keeping genes, transcription starts at the 'initiator' site (CTCANTCT) overlapping the 3' end of the promoter and the 5' end of the first exon.[18] In other cases, more than one TATA box is functional in the promoter region resulting in a variable length of the 5' untranslated region of the mRNA.[19] The TATA binding complex probably starts by the binding of the TATA-box binding protein (TBP) and the interaction of TBP with TBP-binding factors. Some of these factors are essential for transcription and form the so-called minimal initiation complex TFIID. This minimal complex is capable of basal transcription but lacks the ability to cross-talk with other transcription factors. [20] For this function a complete initiation complex encompassing up to 20 proteins is required, including additional basal transcription factors (TFIIB, TFIIA, TFIIE, TFIIF, TFIIH) and TBP associated factors (TAFs) (review [21]). It has been shown *in vitro* that the basal transcription factors and TAFs contributing to this complete initiating complex interact with unrelated transcription factors. However, the importance of the different TAFs *in vivo* remains to be assessed (review[22]). To our knowledge no tissue specific TAFs have been identified yet (B, Table 1).

The level of transcription can be modified by transcription factors binding to conserved DNA elements other than the TATA box. The 5' region immediately upstream of the transcription start site is referred to as the promoter. For a number of muscle specific genes it has been shown that DNA-binding sites in the proximal promoter are sufficient to confer tissue specific expression of a selected gene.[17,23] In other cases, clusters of DNA regulatory elements, referred to as enhancers, are located 5' of the promoter, sometimes at a considerable distance.[24] In some genes, the enhancer is located in the first intron.[25] The distinction between promoter and enhancer is rather arbitrary. Generally speaking, the function of DNA regulatory sequences in the promoter region is highly dependent on their orientation and position relative to the TATA box. In contrast, enhancers function equally well in their orientation and at a variable distance from the TATA box.[26] Experimental studies have shown that enhancer elements can even be effective when placed on the 3'-end of a gene.[27]

Transcription factors play a pivotal role in regulating gene expression. These factors form a class of nuclear proteins that bind to specific regulatory DNA sequences (*cis*-elements) and thereby stimulate (or sometimes inhibit) transcription. These proteins have a DNA binding domain and a *trans*-activation domain. The DNA binding domain can encompass an amino acid sequence necessary for dimerization with identical (homo-) or related factors (hetero-dimers), adjacent to the actual DNA-binding domain. The *trans*-activation domain is physically separated and enables the protein-DNA complex to form a regulatory network in which several factors interact to regulate gene transcription. The recurrent structural motifs in DNA-binding domains provide the basis for the classification of transcription factors: helix-loop-helix, leucine zipper, helix-turn-helix (homeodomain), and the zinc fingers. The common feature of these domains is charged amino acids flanking a helical structure. Most characterized transcription factors share homology in these domains; however, other DNA binding domains have been described like the ETS domain[28], the Fork-Head domain[29], the RNA recognition motif[30] and the MADS-box binding domains.[31] Several different protein domains responsible for *trans*-activation have been characterized. The Glucocorticoid Hormone Receptor has an acidic α-helix domain which stabilizes the transcription initiation complex.[32] The mechanism by which the polar but uncharged glutamine domains found in many transcription factors, like SRF, Sp1 and OCT 1[33], mediate their influence has not been elucidated yet. But the importance of the glutamine domain in Sp1 has been established by deletion analysis.[34] Recent studies have provided support for either a direct interaction of the Sp1 glutamine domain with the TBP in human and *Drosophila*[35] or indirect through co-activators like $dTAF_{11}$ 110. [36] In SRF, an additional proline rich domain has been found to be important for *trans*-activation.[37] The proline gives a fixed kink in a polypeptide and breaks the α-helical structure. Most transcription factors are ubiquitously expressed. However, some factors are expressed in a tissue-restricted or truly tissue-specific fashion (C, Table 1).[38,39] The mechanism by which the binding of transcription factors to their cognate *cis*-elements affects RNA synthesis is still a subject of intensive research. The mechanism has been elucidated for transcription factors that contain a high mobility group domain. These proteins can change the DNA morphology by creating sharp bends. For instance, the leucocyte enhancer binding factor 1 bends the DNA in such a way that a stereo-specific enhancer complex is formed with two additional enhancer binding proteins.[40] It is generally assumed that the transcription factors are capable of bending and folding the DNA in a way more favourable for transcription by RNA polymerase.

Competition and interaction between transcription factors determine the DNA-protein complexes formed at specific sites (D, Table 1). Competition between SRF and YY1 for the same DNA element, the CArG-box or serum responsive element (SRE), has been shown for the skeletal α-actin gene.[41] Alternatively,

proteins can inactivate transcription factors by interfering with the DNA binding domain (ATP-dependent inhibitor[42]). Proteins that lack the actual DNA binding site but can dimerize with other family members can have a negative regulating effect. A good example is the inhibitor of differentiation (Id), that is able to dimerize with other helix-loop-helix (HLH) proteins and prevents them from binding to the E-box site.[43] Other proteins interact with transcription factors and modulate binding of the transcription factor to the DNA-site. For instance, by immunoprecipitation a protein was identified which forms a complex with the transcription factor EFI_A (rat Y-box binding protein, Table 2) on the ventricular myosin light chain 2 (MLC-2v) promoter.[44] These interactions indicate the complexity of the DNA-transcription factor network.

By the process of alternative splicing, different proteins can be derived from the same gene. As discussed later, this type of post-transcriptional modification gives rise to tissue-restricted *Mef*-2 gene products [31] (E, Table 1).

Finally, the function of a transcription factor can be modified through phosphorylation (F, Table 1). This post-translational modification can influence the trans-activation domain of a DNA-protein complex.[45] Other proteins, inactive as monomers, require phosphorylation in order to dimerize and become functional transcription factors.[34, 46]

Strategies to identify transcription factors and functional assays

Identification of transcription factors

Two different approaches can be taken to identify transcription factors. The first approach is to use the DNA-protein interaction as a way to purify the protein or part of the gene encoding the DNA-binding domain. The second strategy is restricted to the DNA level. By using the polymerase chain reaction (PCR) with degenerate primers new related proteins with a conserved domain can be identified.

To purify proteins a method was developed using DNA affinity chromatography. A concatenated DNA molecule containing a repeat of the high-affinity binding site for the transcription factor is linked to activated sepharose support on a column. Nuclear extract harvested from selected tissue or cells in culture can be passed over the column and the purified proteins can be washed from the sepharose afterwards by changing the salt concentration. [47] The amino acid sequence of the purified protein can be determined and translated into DNA sequences. Due to redundancy of the genetic code there will be several different oligonucleotides capable of encoding a particular amino acid sequence. A mixture of the potentially correct oligonucleotides can be synthesized, labeled and used to screen a cDNA library.[37] The purified protein can also be applied to

generate antibodies. An alternative method is to screen a cDNA library for DNA-protein interaction. An expression library has to be constructed in which cDNA is inserted into the coding region of the bacteriophage lambda β-galactosidase gene. Whenever the DNA binding domain of the protein is translated correctly as a part of the resulting bacterial fusionprotein this can bind to a labeled probe. For this purpose a concatenated DNA molecule is synthesized, containing a repeat of the DNA-site involved in the DNA-protein interaction. This method is called Southwestern Screening, as it involves the interaction between protein and DNA.[48] Using the same bacterial fusionprotein, antibodies can be applied to identify clones encoding specific epitopes.[7,33]

If a certain DNA binding domain has been identified, it is possible to look for related clones by designing oligonucleotides with degeneracy in the PCR. The amplified fragment can be cloned, sequenced and used as a probe to screen a cDNA library.[49] With this method different proteins (e.g. herpesvirus gene homologs[50]) with a conserved domain can be cloned.

Functional assays in vitro

Purified protein, nuclear extract, antibodies and DNA molecules can be used to study the DNA-protein interaction in more detail *in vitro*. Methods have been developed to localize the DNA sites involved in DNA-protein interaction and with additional techniques it is possible to identify the single nucleotides essential for protein binding.

DNase I footprinting allows localization of DNA sites. For this method one DNA strand is labeled and incubated with nuclear extract. After DNA-protein binding the DNA can be digested briefly and incompletely by the endonuclease enzyme DNase I. The DNA site bound to protein is protected against the random digestion by DNase I. Analysis of this reaction by electrophoresis on a polyacrylamide gel after removal of the protein will reveal the protected DNA sites.[51]

A free labeled DNA probe will move through a gel at a faster speed compared to a DNA molecule bound to a protein. The shift in mobility during electrophoresis helps to identify the presence of factors binding to the DNA site of interest (gel mobility shift). In this assay specificity of the DNA-protein interaction can be shown by adding an unlabeled competitor. In addition, the influence of mutations in the DNA site can be assessed. If specific antibodies are available against a factor they can be used to show an additional slowing of mobility as the DNA-protein complex is getting bigger (supershift).[52] From this assay information can be derived regarding the size of the factor and a positive supershift helps to characterize the binding protein.

For fine tuning of the DNA protein interaction diethyl pyrocarbonate (DEPC) interference can be used to identify exactly the adenosine and guanine nucleotides involved in protein binding. DEPC reacts with purines bound to protein

generating an alkali-labile abasic site. For this assay single strand DNA is labeled and incubated with a nuclear extract. The DNA-protein complex is treated with DEPC and thereafter the DNA is cut by piperidine at the modified purine nucleotides. The nucleotides bound to protein are recognized by a strong signal after autoradiography of the gel. A sequencing reaction of the same DNA molecule serves as a size marker of this reaction.[33] This method of methylation interference has been modified to be applicable to genomic DNA. The genomic DNA of cultured cells can be treated by dimethylsulfate and again the cleavage sites of the methylated DNA can be assessed.[52,53]

Furthermore the functional importance of selected DNA-sites can be assessed by using reporter genes under the control of wild type or mutated DNA sites. The constructs can be introduced into cells in culture and after an incubation period the activity of the reporter gene can be measured. For these studies neonatal rat cardiomyocytes are cultured and the foreign plasmid DNA can be introduced in the myocytes by direct injection or the calcium phosphate precipitate method.[54, 55] Data derived from these experiments form the bases for the construction of transgenes to further establish the role of identified DNA-sites *in vivo* in the intact transgenic organism.

In vitro model for cardiogenesis

The possibility to use transgenic and gene targeting technology to elucidate the very early steps in murine differentiation are restricted *in vivo*. Knock out experiments resulting in early embryonic death are difficult to analyze. Furthermore the available cell and molecular material is limited. During the last decade alternative models for cardiogenesis have been described and validated.[56] Murine embryonic stem (ES) cells can multiply in culture and the undifferentiated state can be maintained by growing these cells on a feeder layer of fibroblasts or by adding the leukemia inhibitory factor (LIF). In the absence of LIF the ES cells will continue to grow, form clusters, and start to differentiate into so called embryoid bodies. Embryoid bodies derived from the D3 ES cell line will exhibit spontaneous contractions after several days in culture. mRNA analysis of early stages up to day 16 showed the initiation of the cardiac gene program, while no skeletal muscle genes could be identified. Despite the absence of positional cues a marker for chamber specification the ventricular myosin light chain 2 (MLC-2v) is expressed in a time dependent mode similar to normal *in vivo* development. This MLC 2 isoform is specific for cardiac ventricular cells throughout cardiogenesis.[57] The model was validated furthermore by demonstrating the presence of identical ion channels in isolated cells from these bodies and embryonic cardiomyocytes.[6] As the ES cells can be genetically manipulated this *in vitro* model of cardiogenesis should allow studies on ligand dependent transcription and the effect of transcription factor blockade or overexpression on the cardiac

gene program. Alternatively the same modified ES cell line can be used to generate transgenic or gene targeted mice.[5]

Functional assays in vivo

Transgenic and gene targeting technology can also be used to study the specific role of a transcription factor in the complete organism. Basically it is possible to over-express a protein in transgenics by making a constitutively active form (dominant positive) of the protein, or by using an established active promoter linked to the coding region of the gene. A tissue specific promoter will provide over-expression predominantly in the target tissue.[58,59] To determine the effect of reduced gene expression a modified protein can be used that interferes with the function of the target gene (dominant negative) or a knock-out experiment can attenuate gene expression in a homozygous organism.[60] Recent progress in this technology provided a system for tissue specific knock-out experiments.[61,62]

Cis-elements, trans-acting factors and the cardiac muscle phenotype

As more regulatory regions of muscle genes are being characterized and new transcription factors are being identified, it is becoming apparent that tissue-specific gene expression can be brought about in various ways. After the discovery of the skeletal muscle determination factor, MyoD, there was a surge of optimism that soon thereafter, similar factors would be identified that would play a dominant role in determining the phenotype of other muscle cell types, i.e. the smooth muscle and the cardiac muscle cell. As will be shown for the cardiac cell, this search did not lead to the discovery of such a "master transcriptional regulator". Cell fusion studies with cardiomyocyte/fibroblast heterokaryons provided additional evidence that the cardiomyocyte lacks dominant regulatory factors. Contrary to the skeletal muscle/fibroblast heterokaryons, cardiomyocyte/fibroblast heterokaryons do not express muscle specific genes, like *Mef*-2.[63]

Current knowledge suggests that the interaction between transcription factors, either ubiquitous or tissue restricted, and/or their post-transcriptional modification might be the key to understanding the cardiac muscle phenotype (C, D, E, Table 1). In this section the various *cis*-elements and their cognate *trans*-acting factors, when identified, that have been shown to be involved in regulating cardiac gene expression, will be reviewed (Table 2).

The E-box and the bHLH proteins

The discovery of the dominant skeletal muscle gene programme, induced by the myogenic factors MyoD, Myogenin, Myf-5 and MRF4 [64,65] led to the expectation that similar basic HLH proteins might be involved in determining the cardiac gene programme. The basic HLH proteins bind as heterodimers to the consensus DNA sequence CANNTG, the so-called E-box (originally referred to as MEF-1 site). The positively charged basic region is responsible for DNA binding, whereas the HLH region is involved in dimerization. The muscle specific myogenic HLH proteins have been shown to dimerize preferentially with ubiquitous HLH proteins, in particular E12 and E47, splice products of the E2A gene.[66,67] HLH proteins that lack the basic DNA binding region have also been cloned. The most important representative of this class of HLH proteins is Id.[43] Dimerization of MyoD with Id gives rise to a non-functional dimer. Likewise, interaction of Id with E12/E47 lowers the availability of dimerization partners for MyoD. Accordingly, the activity of MyoD is determined by the cellular levels of its dimerization partners, E12/E47 acting as enforcers and Id as an inhibitor.

So far, virtually all attempts to identify a cardiac MyoD-homologue have been futile. Only one group identified a bHLH protein that possibly plays a role in cardiac development. Using an antibody against the second helix of the HLH-domain of MyoD, Bader and coworkers[68] detected an antigenically related protein that is transiently expressed in cardiomyocyte nuclei in the developing avian heart. Subsequently, the antibody was used to screen a chick embryo expression library and a cDNA, referred to as CMF1, was isolated and characterized. Although the CMF1 protein was found to act as a bHLH protein (binds E-box, dimerizes with ubiquitous HLH proteins) it showed limited sequence homology with MyoD.[69] Although the ubiquitous HLH proteins E12, E47, and Id are expressed in the heart and do contribute to transcriptional regulation, their role seems to be modest compared to the myogenic factors in skeletal muscle.[70] The cellular level of Id is developmentally regulated and rapidly increases in response to α-adrenergic stimulation of neonatal and adult cardiomyocytes.[71] However, the forced overexpression of Id does not affect the expression of reporter genes under the control of E-box containing promoters of muscle genes such as the muscle creatine kinase (MCK) and MLC-2v gene.[70]

The importance of the E-box in different promoters may differ from cell-type to cell-type. The muscle specific enhancer of the rat β-myosin heavy chain (β-MHC) promoter is not activated by MyoD in cardiac and skeletal myocytes, but is activated by myogenic factors in fibroblasts.[72] In contrast, the α-MHC promoter appears to be E-box dependent in both the skeletal and cardiac context.[73,74]

Table 2. Transcription factors and their DNA recognition site contributing to the myocyte gene program.

Transcription factor	Species	DNA site	Regulated gene	Regulation	Ref.
Helix Loop Helix					
E12/E47	rat	E-box	MLC-2v	-	70
Id	rat	-	skeletal α-Actin	+	71 #
CMF 1	chicken	E-box		?	151
Twist	drosop.	E-box			152
USF	mouse	E-box	MLC-2v	+/-	52
Homeobox					
Tinman	drosop.	?	Bagpipe	+	81, 153
Nkx 2.5*	mouse	?	?	+	39
Oct 1	mouse	A/T rich	MCK	+	154
Mhox	mouse	A/T rich	MCK	+	92
Zinc finger					
Thyroid hormon receptor	mouse	TRE	α MHC	+	106, 155
Retinoic Acid receptor	mouse	RARE	Hoxb-1	+/-	156, 157
Estrogen receptor	rat	AGGTCA	Connexin 43	+	158
HF1b	rat	A/T rich	MLC- 2v	+	118
GATA 4	mouse	GATA element	cTNT, αMHC	+	120, 124
Sp 1	rat	GC-box	skeletal α-Actin	+	130, 159
YY1	chicken	SRE	skeletal α-Actin	-	41, 160
MADS					
SRF	human	CArG	cardiac Actin	+	161
			MLC-1/3	+	138
			Dystrophin	+	139
Mef 2	rat	A/T rich	MCK	+	38, 138
			MLC-1/3	+	162
			cTNT, MHC	+	23
			MLC-2v	+	17
			Myoglobin	+	107
D-mef 2	drosop.	A/T rich	MHC	+	108
SL1	xenopus	A/T rich	xMLC-2	+	163
Fork-Head domain					
MNF	human	CACC-box	Myoglobin	+	143
Undefined					
TEF-1	human	M-CAT	cTNT, βMHC	+	127
			skeletal α-Actin	+	131

After adrenergic stimulation in cell culture. * in vitro assay. Abbreviations: MLC-2v: (ventricular) myosin light chain 2, MCK: muscle creatine kinase, MHC: myosin heavy chain, cTNT: cardiac troponinT, drosop.: drosophila.

Recently, another ubiquitous bHLH protein, the upstream stimulatory factor (USF), was found to modulate the transcriptional activity of the ventricular MLC-2 promoter and possibly also of the α-MHC promoter.[52]
Murine knock-out experiments targeting MyoD or myf-5 revealed the functional redundancy of the myogenic factors, as each of these homozygous null (-/-) mutants had normal skeletal muscle development. However, a combined knock-out of MyoD (-/-) and myf-5 (-/-) showed a complete absence of skeletal muscle development. It is noteworthy that in these mice, cardiac development appeared to be unaffected.[75] Myogenenin (-/-) mice lack differentiated skeletal muscle but the precursor myoblast population was present and the heart appeared to be normal.[76] Knock-out of the E2A gene, encoding the ubiquitous bHLH-partners E12/E47 had no effect on the differentiation of cardiac muscle cells in the embryonic stem-cell derived embryoid bodies.[77]

Homeobox proteins

Analysis of *Drosophila* mutants allowed the cloning of genes involved in body patterning.[78] Through this work a set of genes was identified involved in either caudo-cranial or lateral patterning. The proteins have a conserved domain in common, the homeodomain, and act as transcription factors. The homeodomain is a conserved 61 amino-acid DNA-binding motif, containing a helix-turn-helix (HTH) domain. Some of the homeobox genes are expressed in a tissue restricted manner. They have multiple functions and play an important role in cell migration and normal embryonic development. The importance of the homeobox proteins for cell migration was shown by injection of goosecoid homeobox mRNA into *Xenopus* embryonic cells, where a region specific cell migration could be induced.[79] A crucial role for these proteins during development is supported by their spatial and temporal expression patterns in, for example, the heart. In addition, murine homeobox gene-targeting experiments exhibited cardiac malformations and were sometimes complicated by early embryonic death due to cardiac abnormalities. The development of a complex organ as the heart, with the formation of valves and the integration of the conduction system, requires precise interaction of cell growth, migration and differentiation. The different homeobox proteins expressed in the heart are believed to be key factors during the process of cardiogenesis. The homeobox gene Tinman proved to be crucial for the development of the dorsal vessel in *Drosophila*.[80,81] The dorsal vessel is considered to be the cardiac homologue in the fly. Using oligonucleotides with sequences derived from this gene, its mammalian homologue (Nkx 2.5, Csx) was cloned from genomic mouse DNA. Nkx 2.5/Csx is expressed predominantly in the cardiac progenitor cells early in development (7.5 dpc). In the adult mouse, the mRNA is present in the heart [39,82] and, albeit at much lower levels, also in spleen and stomach. Knock-out experiments

proved the importance of the gene for cardiogenesis as the homozygous (-/-) embryos died in utero (9 dpc) with a persistent linear heart tube.[83] In the mutant hearts, ventricular markers like β-MHC and cyclin DII were detectable by in-situ hybridization. However, no ventricular MLC-2 expression was detected.[83] The absence of this marker for ventricular chamber specification and the phenotype of the (-/-) embryos is suggestive for an important role for Nkx 2.5 in the process of regionalization within the heart tube. In *Xenopus* an additional homeobox gene was identified (xNkx-2.3) with an expression pattern similar to *Xenopus* Nkx-2.5, as shown by in-situ hybridization. These findings are suggestive of a family of genes representing the *Drosophila tinman* gene in vertebrates.[84]
In chicken, other homeobox genes were found to be associated with valve formation and septation (*Msx*-1), and in the formation of the cardiac conduction system (Msx-2).[85]

Insertional disruption of various homeobox genes resulted in neonatal death[86,87] or a cardiac malformation.[60,88] The Hox gene family is expressed in the heart. Hoxa-3 (-/-) mice display abnormalities resembling the DiGeorge syndrome described in man. The malformations observed include thickening of the ventricular wall, atrial dilatation, defective aortic and pulmonary valves and a persistent ductus arteriosus.[60] The Pax-3 gene is a homeobox gene involved in lateral pattern formation. It is expressed in neural crest cells and the lateral dermomyotome.[89] The (-/-) mice exhibit an attenuated migration of neural crest cells to the cardiac outflow tract and die at 14 dpc due to cardiac abnormalities.[88] Furthermore, the (-/-) mice do not develop limb musculature. Mutations in the homologous human gene (PAX-3) are responsible for an autosomal dominant inherited disorder: the Waardenburg's syndrome.[90]

Many homeobox genes appear to be expressed in a gradually changing concentration (G, Table 1); another phenomenon that may play a role in developmental regulation of gene expression [91]. As yet, it is unclear which genes are controlled by these homeobox proteins and which DNA-elements are involved in DNA binding. Apparently the sensitivity of different promoters for the presence of these transcription factors is variable. In the case of Mhox, another member of the murine homeobox family, it has been shown that it binds to an A/T rich DNA element on the MCK enhancer. This element is essential for muscle specific gene expression. The Mhox protein is restricted to muscle tissue.[92] *In vitro* studies, using bacterially produced Nkx-2.5 protein and randomly generated oligonucleotides, showed relatively weak binding affinity for A/T rich sites.[93] In this study, the Nkx-2.5 showed highest affinity for the consensus sequence TNAAGTG.

Phox, a human homeobox protein, stimulates the binding of SRF to the SRE. For mediating this process the Phox protein itself does not interact with the DNA. As these proteins are expressed in a tissue-restricted pattern, this potentiating mechanism could evoke a muscle specific response to a non-specific stimulus.[94]

(Steroid) hormone receptor superfamily

A crucial role in normal development is mediated by hormones that exert their influence on transcription through nuclear receptors with the zinc finger motif. This rapidly growing group of the nuclear receptor superfamily has some striking common features (Table 2). Substances like retinoic acid (RA), thyroid hormone, vitamin D3 and steroid hormones pass the plasma membrane easily and bind their receptors within the cell. Nuclear receptors can be found in the cytoplasm and are translocated to the nucleus after ligand binding. Part of the regulation of these receptors therefore may occur in the cytoplasm. The activated receptors bind to promoter sequences of target genes as homo- or heterodimers and, in this way, modify the rate of transcription. One of the common features of the nuclear receptor hormone complex is the recognition of very closely related DNA elements. The ligand-receptor complexes recognize the consensus sequence AGGTCA. For efficient binding of the heterodimer this element has to be present as a direct or inverted repeat often separated by 2, 3, 4 or 5 bases. The spacing of the repeats confers some degree of specificity to the binding of heterodimers formed by individual members of the nuclear receptor family.[95] The most commonly formed complex is a heterodimer of RXR with one of the other nuclear receptors on the repeat motif. Besides these hormone binding transcription factors, a group of homologous proteins has been identified often binding DNA as monomers, without established ligand interaction and variable influence on transcription. Some *in vitro* studies suggest that these 'orphan' receptors may have antagonistic functions within the same cell. One receptor may stimulate transcription while a related factor can compete for the same DNA-binding site and suppress transcription.[96] A specific orphan receptor (RORγ) that is highly enriched in skeletal muscle was recently described.[97]

Retinoic acid (RA) is a vitamin A derivative and is considered to be a morphogen, i.e. a signalling molecule involved in pattern formation and morphogenesis. There is evidence pointing to the importance of a RA gradient in differentiation induction (G, Table 1). In chicken, homeobox genes are induced in response to RA treatment, suggesting that cross-talk between the homeobox and hormone receptor gene families is important during development (review[98]).

The importance of retinoid X receptor-α (RXR-α) for normal cardiac development was recently shown by studying homozygous (-/-) mice.[99-101] RXR-α is one of six retinoic acid receptors: three retinoic acid receptors (RAR) and three retinoid X receptors.[102] Histologic studies in RXR-α(-/-) mouse embryos revealed a normal development up to 11.5 dpc, but the embryos die at 14 dpc due to cardiac defects and generalized oedema, which could be a manifestation of embryonic congestive heart failure. The heart at this stage displayed a ventricular septal defect and ventricular hypoplasia. A method was developed to assess left ventricular function in these embryos and revealed a significant impairment of ejection

fraction of 50% in wild type vs 15% in (-/-) mice.[6] At the transcriptional level, a persistent expression of an atrial specific isoform of MLC-2(a) was found in the (-/-) embryos. During normal cardiac development, MLC-2 a can be used as a marker for chamber specification and is present in ventricular tissue up to 11 dpc.[103] These findings suggest a crucial role of the RXR-α receptor in chamber specification during development of the mammalian heart.

An interesting feature of the regulation of the α-MHC gene is its sensitivity to stimulation with thyroid hormone in ventricular myocytes, but not in atrial cardiomyocytes.[104] After birth, the α-MHC becomes the most predominant isoform in the ventricular myocardium in rodents, while β-MHC expression is down-regulated. The natural increase in circulating thyroid hormone level after birth is primarily responsible for the switch in MHC isoforms.[105] Accordingly, the thyroid hormone responsive α-MHC promoter has been used to drive the expression of transgenes in a developmentally regulated (no transcription before birth, high level of transcription after birth) and tissue specific (ventricular myocytes) manner.[106]

The A/T-rich sequence and MEF-2

MEF-2 (Myocyte Enhancer Factor 2) was originally described as a DNA binding activity found in nuclear extracts of differentiated skeletal muscle cells.[38] MEF-2 binds to A/T-rich elements that are present in the promoter and enhancer regions of many genes expressed in both skeletal and cardiac muscle (Table 2).[17,23,38]

By using the A/T-rich consensus sequence as a probe, a large number of cDNAs that possess MEF-2 activity have been cloned from muscle expression libraries. The MEF-2 proteins appear to be splice variants of four genes, designated Mef2a, Mef2b, Mef2c, and Mef2d.[31] The MEF-2 proteins contain the so-called MADS-box, a conserved sequence of 56 amino acids, responsible for both DNA-binding and dimerization, which is also found in, for instance, the DNA-binding protein SRF; hence, the synonym RSRF (related to serum response factor) for MEF-2 proteins.[107] One characteristic of all MEF-2 proteins is an additional conserved 29 amino acid sequence adjacent to the MADS-box.
In *Drosophila* only one MEF-2 gene (D-mef2) is present. In loss-of-function type of experiments, the inactivation of this gene resulted in the inability of the cells to differentiate into any of the myogenic lineages, i.e. the somatic, cardiac, or visceral muscle lineage.[108,109] Conversely, in gain-of-function experiments the forced overexpression of the Mef2a gene in cultured non-muscle cells, like 10T1/2 fibroblasts, converted these cells to the skeletal muscle lineage.[110] This response is reminiscent of the effect of the bHLH protein MyoD on permissive cell lines.[111] Both the loss-of-function and gain-of-function studies emphasize the important role of MEF-2 proteins in muscle development.

However, research designed to elucidate the role of the Mef2 gene products is not as straightforward as the experiments described above suggest. The pattern of expression of splice variants of the four genes is highly complex. In general, studies at the RNA level indicate that the transcripts are ubiquitously expressed, although expression levels are generally higher in muscle cell types. Paradoxically, at the protein level, MEF-2 appears to be restricted to a limited number of tissues, among which are muscle and brain. [31,38,112,113] Indeed, the presence of transcripts in certain cell types does not always coincide with the presence of MEF-2 protein and DNA-binding activity.[31] Contrasting results, however, were reported by Pollock and Treisman and Chambers et al[107,114], who showed that MEF-2 binding activity is ubiquitous.

It is becoming apparent that alternative exon usage plays an important role in determining tissue specificity. Indeed, the existence of muscle- and brain-specific exons has been demonstrated.[115,116] These observations also support the idea that, in the case of MEF-2, post-transcriptional events are important in bringing about tissue specific gene expression (E, Table 1). Interestingly, Mef2c mRNAs are very abundant in skeletal muscle, brain and spleen, but absent in heart tissue.[113,116,117] Theoretically, dimerization of all splice variants of the four genes could give rise to over a hundred different heterodimers. It has been argued that this could provide an effective means to establish combinatorial control of gene expression.[115]

Finally, it should be noted that the A/T-rich sequence has also been found to bind factors other than MEF-2. The A/T-rich region in the proximal promoter of the MLC-2v gene is able to bind a tissue-restricted zinc finger protein, HF1-B. Overexpression of this zinc finger protein upregulated the expression of co-transfected MLC-2 promoter/reporter hybrid genes.[118] Likewise, the A/T-rich element of the myoglobin promoter binds both MEF-2 and an unrelated protein referred to as ATF35.[69] Possible explanations could be that the flanking regions of the A/T-rich element affect binding specificity and/or that MEF-2 and the other binding factors bind to overlapping sites. Alternatively, dimerization between MEF-2 and other factors may play a role.

The GATA element and zinc finger proteins

The consensus DNA sequence (A/T)GATA(A/G), referred to as the GATA element, has been found to be an important *cis*-element in the regulation of a variety of genes. Several proteins binding to the GATA element have been cloned. The GATA-binding proteins (GATA-1 to -4) are characterized by two DNA-binding zinc finger domains. One member of this family of transcription factors, GATA-4 has been implicated in cardiac differentiation.[119,120] Its expression is restricted to the heart, intestinal epithelium, primitive endoderm, and gonads. [121-123] GATA-4 is not detectable in skeletal muscle. In embryoid bodies, the

increase in GATA-4 mRNA paralleled the development of spontaneous contractile activity. In addition, GATA-4 expression appeared to be RA responsive [121], which is further supported by the observation that knock-out of RAR-γ abolished GATA-4 induction.[99]

GATA-4 has been implicated in the transcriptional regulation of slow cardiac troponin C (cTnT), brain natriuretic factor, and α-MHC.[120,124,125] α-MHC is the predominant isoform in the rodent ventricle and is absent in skeletal muscle.[119,120] α-MHC contains two putative GATA elements in an enhancer region near the promoter. Mutation of either of these sites substantially reduced transcriptional activity. Interestingly, overexpression of GATA-4 led to the ectopic expression of a co-transfected α-MHC promoter/enhancer construction in skeletal muscle cells, but not in fibroblasts. This finding suggests that the absence of α-MHC expression in skeletal muscle *in vivo* is due to the absence of GATA-4 in these cells. Altogether, the recent observations regarding GATA-4 might be instrumental in solving the mechanisms behind the alternative gene programs in skeletal and cardiac muscle.

The M-CAT element and TEF-1

The M-CAT element is a *cis*-element that is found in several muscle genes (Table 2). Originally, the sequence specific DNA-binding activity was referred to as MCBF (M-CAT Binding Factor).[126] Using gel mobility shift assays and immunological techniques, this binding activity appeared to be indistinguishable from TEF-1 (Transcriptional Enhancer Factor 1; not to be confused with Translation Elongation Factor 1), a transcription factor of viral enhancers.[127,128] TEF-1 is ubiquitously expressed, but the highest levels are found in skeletal and cardiac muscle.[128,129] In addition to mediating tissue specificity of several muscle genes, TEF-1 also seems to be involved in conferring inducible expression of these genes. Stimulation of neonatal cardiomyocytes with α-adrenergic agonists or transforming growth factor-β enhanced the transcription of the β-MHC and skeletal α-actin promoters.[130-132] Overexpression of a constitutively active form of β-protein kinase C resulted in the enhanced transcription of β-MHC through TEF-1 dependent pathways. [131]

Using an entirely different experimental approach, Chen *et al.*[133] characterized a mutant strain with a recessive lethal embryonic phenotype. In these embryos cardiac malformations were a predominant feature. Subsequent characterization of this mutant strain indicated that the TEF-1 gene was disrupted, suggesting a causal relationship between TEF-1 expression and cardiogenesis. All these observations point to an important role of TEF-1 in cardiac gene expression. However, in at least one study the relevance of TEF-1 has been questioned. In this study transgene technology was applied to study the role of the M-CAT site within the β-MHC promoter *in vivo*. Surprisingly, mutagenesis

of the M-CAT site did not affect the level and tissue distribution of the expression of the chloramphenicol acetyltransferase reporter gene, placed downstream of the β-MHC promoter.[134]

The CArG motif and serum response factor

The CArG box [$CC(A/T)_6GG$] is another A/T-rich type of *cis*-element important in regulating the transcription of the skeletal, cardiac, and smooth muscle α-actin genes[73,130,135-137], MLC-1/3[138], and the Duchenne muscular dystrophy gene[139] (Table 2). Evidence is accumulating that the CArG-box is rather promiscuous in the binding of *trans*-acting factors. The CArG box is identical to the SRE found in several immediate early genes, including *c-fos*. DNA-binding studies revealed that the ubiquitous MADS-box protein SRF binds to the CArG/SRE element of both types of promoter.[73,140] Another protein, distinct from SRF, was also able to interact with the CArG-box of the skeletal α-actin gene. This activity was originally referred to as F-ACT1 and later found to be identical to the transcription factor YY1.[41] The binding of SRF and YY1 appears to be mutually exclusive, YY1 acting as a negative regulator. Furthermore, the ANF promoter also contains a CArG-box motif that was shown to interact with a factor, unrelated to SRF. This factor behaves like a cardiac specific binding complex.[141] Recent findings suggest that the homeobox proteins Nkx2.5/Csx might interact with the core region of the CArG-box.[39] Again, the study of the CArG-box motif indicates that cross-talk between different transcription factors, in this case competition for the same *cis*-element, is an important mechanism in transcriptional control.

Complexity of the regulatory network

It should be realized that this enumeration of *cis*-elements and *trans*-acting factors relevant to the cardiac gene programme is not exhaustive. In the upstream promoter region of the myoglobin gene, other sequence elements, in addition to the MEF-2 site, partake in transcriptional control. A so-called CACC box is essential for skeletal and cardiac muscle expression.[142] The subsequent cloning of the binding activity demonstrated that the factor, designated MNF (Myocyte Nuclear Factor), is related to *fork head* and hepatocyte nuclear factor 3a, members of the winged-helix family of transcription factors.[142,143] Moreover the detailed characterization of the ANF and cardiac troponin T (cTnT) promoters revealed DNA-binding activities that are unique for cardiac tissue.[69,141,144]

The latter findings suggest that still other cardiac specific transcription factors are likely to be cloned in the near future. Many muscle protein isoforms are expressed in skeletal as well as cardiac muscle cells. Certain promoter/enhancer regions are required for efficient expression of those genes in both cell types. Often, separate regions, in addition to the shared regions, are needed to confer

skeletal muscle or cardiac specificity to these genes. This has been shown, amongst others, for the MLC-2v, MCK, cTnC and cTnT genes [23,25,58,144]. In the case of cTnT, the proximal 129 bp of the promoter, containing two M-CAT elements and one CArG box, is sufficient for skeletal muscle specific expression. This region is required, but not sufficient, for cardiac expression. To confer cardiac specificity an additional region of the promoter, extending from -269 to -201 bp relative to the transcription start site, is necessary.[23] For the MCK gene the situation is quite different. In that case, the proximal promoter is sufficient to drive expression in cardiac muscle cells, whereas additional enhancer elements are required for its expression in skeletal muscle.[144] The cTnT is expressed in heart and slow skeletal muscle in the mouse. For skeletal muscle an element in the first intron is essential. However, the regulation of cardiac expression is ruled by the 5' flanking region of the gene.[25] In humans, both the 5' and the intron element contribute to cardiac muscle specific expression.[4]

As mentioned earlier, tissue specific expression often requires a close interaction between ubiquitous and tissue specific factors. This is probably best illustrated by the identification of the different factors binding to the 28 bp HF-1 region in the proximal promoter of the MLC-2v gene. This region is functionally divided into a HF-1a element, binding the ubiquitously expressed factor EF1A, and an HF-1b element, which is the binding site for the tissue specific factors Mef-2 and HF-1b binding protein.[145, 146] Both sites, HF-1a and HF-1b, are required for the maintenance of cardiac specific expression. A comparable conjunction of DNA binding sites is also found on the α-MHC promoter, where a site resembling HF-1a, is adjacent to a MEF-2 binding site.[147, 148]

Additional studies in transgenic animals suggest the presence of distinctive *cis*-elements and *trans*-acting factors involved in cell specificity within cardiac compartments. The expression of the reporter gene LacZ under control of a 1 kb fragment of the human desmin promoter is confined to skeletal myocytes and a subset of ventricular cardiomyocytes representing the developing conduction system (Li et al, 1993).[59] Surprising results were obtained with the fast skeletal muscle MLC-3F promoter in conjunction with a muscle specific enhancer located 3' of the coding region of the MLC-3F gene. In this case, LacZ expression was observed in the atria and left ventricle, but not in the right ventricle. The latter finding suggests that there are differences in transcriptional regulation between the left and right ventricle of the heart.[149] These examples illustrate that during evolution a large variety of 'solutions' have been attained that mediate cardiac specific expression.

Future directions

Different molecular mechanisms contribute to the cardiac muscle phenotype. Although we learned a lot in the past decade, we are still far from understanding the cardiac gene programme. Several of the *trans*-acting factors binding directly to cardiac genes have been identified. Maybe these factors form only the edges of a complex puzzle. Interactions have been shown between hormone and morphogen receptors, and different transcription factors, as part of an extensive regulatory network. The studies discussed here deal mainly with the contractile protein programme, but other genes encoding ion channels, cytoskeletal proteins, adhesion molecules and proteins involved in cardiomyocyte metabolism may appear to be fruitful models for the characterization of new transcription factors and in helping to elucidate the patterns of cross-talk. The invertebrate model (*Drosophila*) has been proven useful in studying the early genetic steps involved in cardiogenesis. The identified genes can be used to clone the evolutionary conserved mammalian counterparts. The murine embryonic stem-cell derived embryoid body system looks promising as an alternative source of cardiac precursor cells and provides a model to study cellular commitment and differentiation during cardiogenesis. In addition, the stem cells can be manipulated relatively easy, allowing gene-targeting and transgenic techniques to evaluate identified and new factors. The embryoid bodies and transgenic organisms can be studied to assess the importance of transcription factors for cardiovascular development and physiology.

Literature

1. Chien KR, Zhu H, Knowlton KU, Miller-Hance W, van-Bilsen M, O'Brien TX, Evans SM: Transcriptional regulation during cardiac growth and development. *Annu Rev Physiol* 1993;55:77-95
2. Olson EN: Regulation of muscle transcription by the MyoD family; the heart of the matter. *Circ Res* 1993;72:1-6
3. van Bilsen M, Chien KR: Growth and hypertrophy of the heart: towards an understanding of cardiac specific and inducible gene expression. *Cardiovasc Res* 1993;27:1140-1149
4. Sartorelli V, Kurabayashi M, Kedes L: Muscle specific gene expression: a comparison of cardiac and skeletal muscle transcription strategies. *Circ Res* 1993;72:925-931
5. Dzau VJ, Gibbons GH, Koblika BK, Lawn RM, Pratt RE: Genetic models of human cardiovascular disease. *Circulation* 1995;91:521-531
6. Kubalak S, Doevendans PA, Rockman H, Hunter JJ, Tanaka N, Ross J, Chien KR: Molecular analysis of cardiac muscle diseases via mouse gentics, in Adolph KW (ed): *Methods in molecular genetics*. Orlando, Academic Press, 1996, pp. 470-487.

7. DeAngelo DJ, DeFalco J, Childs G: Purification and characterization of the stage-specific embryonic enhancer-binding protein SSAP-1. *Mol Cell Biol* 1993;13:1746-1758
8. Segers A, Muyldermans S, Wyns L: The interaction of histone H5 and its globular domain with core particles, depleted chromatosomes, polynucleosomes, and a DNA decamer. *J Biol Chem* 1991;266:1502-1508
9. Bresnick EH, Bustin M, Marsaud V, Richard Foy H, Hager GL: The transcriptionally-active MMTV promoter is depleted of histone H1. *Nucleic Acids Res* 1992;20:273-278
10. Dimitrov.S, Wolffe.A: Chromatin and nuclear assembly: experimental approaches towards the reconstitution of transcriptionally active and silent states. *Biochimica et Biophysica Acta* 1995;1260:1-13
11. Wolffe AP: Transcription: in tune with the histones. *Cell* 1994;77:13-16
12. Roth SY, Shimizu M, Johnson L, Grunstein M, Simpson RT: Stable nucleosome positioning and complete repression by the yeast alpha 2 repressor are disrupted by amino-terminal mutations in histone H4. *Genes Dev* 1992;6:411-425
13. Paroush Z, Keshet I, Yisraeli J, Cedar H: Dynamics of demethylation and activation of the α-actin gene in myoblasts. *Cell* 1990;63:1229-1237
14. Meehan RR, Lewis JD, McKay S, Kleiner EL, Bird AP: Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* 1989;58:499-507
15. Li E, Bestor TH, Jaenisch R: Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 1992;69:915-926
16. Bird A: The essentials of DNA methylation. *Cell* 1992;70:5-8
17. Zhu H, Garcia AV, Ross RS, Evans SM, Chien KR: A conserved 28-base-pair element (HF-1) in the rat cardiac myosin light-chain-2 gene confers cardiac-specific and alpha-adrenergic-inducible expression in cultured neonatal rat myocardial cells. *Mol Cell Biol* 1991;11:2273-2281
18. Smale ST, Schmidt MC, Berk AJ, Baltimore D: Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID. *Proc Natl Acad Sci U S A* 1990;87:4509-4513
19. Simeone A, Pannese M, Acampora D, D'Esposito M, Boncinelli E: At least three human homeoboxes on chromosome 12 belong to the same transcription unit. *Nucleic Acids Res* 1988;16:5379-5390
20. Goodrich JA, Tjian R: Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell* 1994;77:145-156
21. Buratowski S: The basics of basal transcription by RNA Polymerase II. *Cell* 1994;77:1-3
22. Tjian R, Maniatis T: Transcriptional activation: a complex puzzle with few easy pieces. *Cell* 1994;77:5-8
23. Iannello RC, Mar JH, Ordahl CP: Characterization of a promoter element required for transcription in myocardial cells. *J Biol Chem* 1991;266:3309-3316
24. Herbst RS, Friedman N, Darnell JE, Babiss LE: Positve and negative regulatory elements in the mouse albumin enhancer. *Mol Cell Biol* 1989;10:3896-3905

25. Parmacek MS, Vora AJ, Shen T, Barr E, Jung F, Leiden JM: Identification and characterization of a cardiac-specific transcriptional regulatory element in the slow/cardiac troponin C gene. *Mol Cell Biol* 1992;12:1967-1976
26. Li Z, Paulin D: Different factors interact with myoblast-specific and myotube-specific enhancer regions of the human desmin gene. *J Biol Chem* 1993;268:10403-10415
27. Semenza GL, Roth PH, Fang HM, Wang GL: Transcriptional regulation of genes encoding glycolytic enzymes by hypoxia-inducible factor 1. *J Biol Chem* 1994;269:23757-23763
28. Karim FD, Urness LD, Thummel CS, Klemsz MJ, McKercher SR, Celada A, Van Beveren C, Maki RA, Gunther CV, Nye JA, et al: The ETS-domain: a new DNA-binding motif that recognizes a purine-rich core DNA sequence [letter]. *Genes Dev* 1990;4:1451-1453
29. Weigel D, Jackle H: The fork head domain: a novel DNA binding motif of eukaryotic transcription factors? [letter]. *Cell* 1990;63:455-456
30. DeAngelo DJ, DeFalco J, Rybacki L, Childs G: The embryonic enhancer-binding protein SSAP contains a novel DNA-binding domain which has homology to several RNA-binding proteins. *Mol Cell Biol* 1995;15:1254-1264
31. Yu YT, Breitbart RE, Smoot LB, Lee Y, Mahdavi V, Nadal Ginard B: Human myocyte-specific enhancer factor 2 comprises a group of tissue-restricted MADS box transcription factors. *Genes Dev* 1992;6:1783-1798
32. Hollenberg SM, Evans RM: Multiple and cooperative *Trans*-activation domains of the glucocorticoid receptor. *Cell* 1988;55:899-906
33. Sturm RA, Das G, Herr W: The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. *Genes & Development* 1988;2:1582-1599
34. Mitchell PA, Tjian R: Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* 1989;245:371-378
35. Emili A, Greenblatt J, Ingles CJ: Species-specific interaction of the glutamine-rich activation domains of Sp1 with the TATA box-binding protein. *Mol Cell Biol* 1994;14:1582-1593
36. Gill G, Pascal E, Tseng ZH, Tjian R: A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the dTAFII110 component of the Drosophila TFIID complex and mediates transcriptional activation. *Proc Natl Acad Sci U S A* 1994;91:192-196
37. Norman C, Runswick M, Pollock R: Isolation and properties of cDNA clones encoding SRF, a transcription factor that binds to the c-fos serum response element. *Cell* 1988;55:989-1003
38. Gosset LA, Kelvin DJ, Strenberg EA, Olson EN: A new myocyte specific enhancer-binding factor that recognizes a conserved element associated with multiple muscle-specific genes. *Mol Cell Biol* 1989;9:5022-5033
39. Lints TJ, Parsons LM, Hartley L, Lyons I, Harvey RP: Nkx-2.5: a novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants. *Development* 1993;119:419-431

40. Giese K, Cox J, Grosschedl R: The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell* 1992;69:185-195
41. Lee T, Shi Y, Schwartz R: Displacement of BrdUrd-induced YY1 by serum response factor activates skeletal α-actin transcription in embryonic myoblasts. *Proc Natl Acad Sci U S A* 1992;89:9814-9818
42. Auble DT, Hahn S: An ATP-dependent inhibitor of TBP binding to DNA. *Genes Dev* 1993;7:844-856
43. Benezra R, Davis RL, Lockshon D, Turner DL, Weintraub H: The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell* 1990;61:49-59
44. Zhou MD, Goswami SK, Martin ME, Siddiqui MA: A new serum-responsive, cardiac tissue-specific transcription factor that recognizes the MEF-2 site in the myosin light chain-2 promoter. *Mol Cell Biol* 1993;13:1222-1231
45. Sorger PK, Lewis MJ, Pelham HR: Heat shock factor is regulated differently in yeast and Hela cells. *Nature* 1987;329:81-84
46. Gonzalez GA, Montminy MR: Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. *Cell* 1989;59:675-680
47. Kadonga JT, Tjian R: Affinity purification of sequence-specific DNA binding proteins. *Proc Natl Acad Sci U S A* 1986;83:5889-5893
48. Vinson CR, LeMarco K, Johnson P, Landschulz WH, McKnight SL: In situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. *Genes Dev 1988;2:801-806*
49. Compton T: Degenerate primers for DNA amplification, in *PCR protocols: A guide to methods and applications.* 1990, pp 39-45
50. Nunberg JH, Wright DK, Cole GE, Petrovski EA, Post LE, Compton T, Gilbert JH: Identification of the thymidine kinase gene of feline herpesvirus: use of degenerate oligonucleotides in the polymerase chain reaction to isolate herpesvirus gene homologs. *J Virol* 1989;63:3240-3249
51. Schmitz A, Galas DJ: DNase I footprinting: a simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res* 1978;5:3157-3170
52. Navankasattusas S, Sawadogo M, van Bilsen M, Dang CV, Chien KR: The basic Helix-Loop-Helix protein upstream stimulating factor regulates the cardiac ventricular myosin light-chain 2 gene via independent *cis* regulatory elements. *Mol Cell Biol* 1994;14:7331-7339
53. Strauss EC, Andrews NC, Higgs DR, Orkin SH: In vivo footprinting of the human alpha-globin locus upstream regulatory element by guanine and adenine ligation-mediated polymerase chain reaction. *Mol Cell Biol* 1992;12:2135-2142
54. Simpson P, Shoshana S: Differentiation of rat myocytes in single cell cultures with and without proliferating nonmyocardial cells. *Circ Res* 1982;50:101-116
55. Fuller SJ, Chien KR: Genetic engineering of cardiac muscle cells: in vitro and in vivo. *Genet Eng N Y* 194;16:17-31.

56. Doetschman TC, Eistetter H, Katz M, Schmidt W, Kemler R: The in vitro development of blastocyst-derive embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J Embryol Exp Morphol* 1985;87:27-45
57. Miller-Hance WC, LaCorbiere M, Fuller SJ, Evans SM, Lyons G, Schmidt C, Robbins J, Chien KR: In vitro chamber specification during embryonic stem cell cardiogenesis. Expression of the ventricular myosin light chain-2 gene is independent of heart tube formation. *J Biol Chem* 1993;268:25244-25252
58. Lee KJ, Hickey R, Zhu H, Chien KR: Positive regulatory elements (HF-1a and HF-1b) and a novel negative regulatory element (HF-3) mediate ventricular muscle-specific expression of myosin light-chain 2-luciferase fusion genes in transgenic mice. *Mol Cell Biol* 1994;14:1220-1229
59. Li Z, Marchand P, Humbert J, Babinet C, Paulin D: Desmin sequence elements regulating skeletal muscle-specific expression in transgenic mice. *Development* 1993;117:947-959
60. Chisaka O, Capecchi MR: Regionally restricted developmental defects resulting from targeted disruption of the mouse homeobox gene hox-1.5. *Nature* 1991;350:473-479
61. Gu H, Marth JD, Orban PC, Mossmann H, Rajewsky K: Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. *Science* 1994;265:103-106
62. Doevendans PA, Hunter JJ, Lembo G, Wollert KC, Chien KR: Strategies for studying cardiovascular diseases in transgenic mice and gene-targeted mice, in Monastersky GM, Robl JM (eds): *Strategies in transgenic animal science.* Washington, American Society For Microbiology, 1995, pp 107-144.
63. Evans SM, Tai LJ, Tan VP, Newton CB, Chien KR: Heterokaryons of cardiac myocytes and fibroblasts reveal the lack of dominance of the cardiac muscle phenotype. *Mol Cell Biol* 1994;14:4269-4279
64. Olsen EN: MyoD family: a paradigm for development. *Genes Dev* 1990;4:1454-1461
65. Weintraub H, Davis R, Tapscott S, Thayer M, Krause M, Benezra R, Blackwell T, Rupp R, Hollenberg S, Huang Y, Lassar AB: The MyoD family: nodal point during specification of the muscle cell lineage. *Science* 1989;251:761-766
66. Blackwell TK, Weintraub H: Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* 1990;250:1104-1110
67. Buskin JN, Hauschka SD: Identification of a myocyte nuclear factor that binds to the muscle-specific enhancer of the mouse muscle creatine kinase gene. *Mol Cell Biol* 1989;9:2627-2640
68. Litvin J, Montgomery MO, Goldhamer DJ, Emerson CP, Jr., Bader DM: Identification of DNA-binding protein(s) in the developing heart. *Dev Biol* 1993;156:409-417
69. Grayson J, Williams RS, Yu YT, Bassel Duby R: Synergistic interactions between heterologous upstream activation elements and specific TATA sequences in a muscle-specific promoter. *Mol Cell Biol* 1995;15:1870-1878.

70. Evans SM, Walsh BA, Newton CB, Thorburn JS, Gardner PD, van Bilsen M: Potential role of helix loop helix proteins in cardiac gene expression. *Circ Res* 1993;73:569-578
71. Springhorn JP, Ellingsen O, Berger HJ, Kelly RA, Smith TW: Transcriptional regulation in cardiac muscle. Coordinate expression of Id with a neonatal phenotype during development and following a hypertrophic stimulus in adult rat ventricular myocytes in vitro. *J Biol Chem* 1992;267:14360-14365
72. Thompson WR, Nadal Ginard B, Mahdavi V: A MyoD1-independent muscle specific enhancer controls the expression of the β myosin heavy chain gene in skeletal and and cardiac muscle cells. *J Biol Chem* 1991;266:22678-22688
73. Sartorelli V, Webster KA, Kedes L: Muscle-specific expression of the cardiac alpha-actin gene requires MyoD1, CArG-box binding factor, and Sp1. *Genes Dev* 1990;4:1811-1822
74. Sartorelli V, Hong NA, Bishopric NH, Kedes L: Myocardial activation of the human cardiac alpha-actin promoter by helix-loop-helix proteins. *Proc Natl Acad Sci U S A* 1992;89:4047-4051
75. Rudnicki MA, Schnegelsberg PN, Stead RH, Braun T, Arnold HH, Jaenisch R: MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell* 1993;75:1351-1359
76. Hasty P, Bradley A, Morris JH, Edmondson DG, Venuti JM, Olsen EN, Klein WH: Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. *Nature* 1993;364:501-506
77. Zhuang Y, Kim CG, Bartelmez S, Cheng P, Groudine M, Weintraub H: Helix-loop-helix transcription factors E12 and E47 are not essential for skeletal or cardiac myogenesis, erythropoiesis, chondrogenesis, or neurogenesis. *Proc Natl Acad Sci USA* 1992;89:12132-12136
78. Rosenberg UB, Schroder C, Preiss A, Kienlin A, Cote S, Riede I, Jackle H: Structural homology of the *drosophila Kruppel* gene *Xenopus* transcription factor III A. *Nature* 1986;319:336-339
79. Niehrs C, Keller R, Cho KW, De Robertis EM: The homeobox gene goosecoid controls cell migration in Xenopus embryos. *Cell* 1993;72:491-503
80. Wu X, Golden K, Bodmer R: Heart development in Drosophila requires the segment polarity gene wingless. *Dev Biol* 1995;169:619-628
81. Bodmer R, Jan LY, Jan YN: A new homeobox-containing gene, msh-2, is transiently expressed early during mesoderm formation of Drosophila *Development* 1990;110:661-669
82. Komuro I, Izumo S: Csx: a murine homeobox-containing gene specifically expressed in the developing heart. *Proc Natl Acad Sci U S A* 1993;90:8145-8149
83. Lyons I, Parsons LM, Hartley L, Li R, Andrews JE, Robb L, Harvey RP: Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene Nkx2-5. *Genes Dev* 1995;9:1654-1666

84. Evans SM, Yan W, Murillo MP, Ponce J, Papalopulu N: tinman, a Drosophila homeobox gene required for heart and visceral mesoderm specification, may be represented by a family of genes in vertebrates: XNkx-2.3, a second vertebrate homologue of tinman. *Development* 1995;121:3889-3899
85. Chan-Thomas PS, Thompson RP, Robert B, Yacoub MH, Barton PJ: Expression of homeobox genes Msx-1 (Hox-7) and Msx-2 (Hox-8) during cardiac development in the chick. *Dev Dyn* 1993;197:203-216
86. Jeannotte L, Lemieux M, Charron J, Poirier F, Robertson EJ: Specification of axial identity in the mouse: role of the Hoxa-5 (Hox1.3) gene. *Genes Dev* 1993;7:2085-2096
87. Rijli FM, Mark M, Lakkaraju S, Dierich A, Dolle P, Chambon P: A homeotic transformation is generated in the rostral branchial region of the head by disruption of Hoxa-2, which acts as a selector gene. *Cell* 1993;75:1333-1349
88. Franz T: Persistent truncus arteriosus in the Splotch mutant mouse. *Anatomy and Embryology* 1989;180:457-464
89. Kern MJ, Argao EA, Potter SS: Homeobox genes and heart development. *Trends Cardiovasc Med* 1995;5:47-54
90. Tassabehji M, Newton VE, Leverton K, Turnbull K, Seemanova E, Kunze J, Sperling K, Strachan T, Read AP: PAX3 gene structure and mutations: close analogies between Waardenburg syndrome and the Splotch mouse. *Hum Mol Genet* 1994;3:1069-1074
91. Driever W, Nusslein-Volhard C: A gradient of bicoid protein in *drosophila* embryos. *Cell* 1988;54:95-104
92. Cserjesi P, Lilly B, Bryson L, Wang Y, Sassoon DA, Olson EN: MHox: a mesodermally restricted homeodomain protein that binds an essential site in the muscle creatine kinase enhancer. *Development* 1992;115:1087-1101
93. Chen CY, Schwartz RJ: Identification of novel DNA binding targets and regulatory domains of a murine tinman homeodomain factor, nkx-2.5. *J Biol Chem* 1995;270:15628-15633
94. Grueneberg DA, Natesan S, Alexandre C, Gilman MZ: Human and Drosophila homeodomain proteins that enhance the DNA-binding activity of serum response factor. *Science* 1992;257:1089-1095
95. Umesono K, Murakami KK, Thompson CC, Evans RM: Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. *Cell* 1991;65:1255-1266
96. Ohno CK, Uedah H, Petkovich M: The *Drosophila* nuclear receptors FTZ-F1α and FTZ-F1β compete as monomers for binding to a site in the *fushi tarazu* gene. *Mol Cell Biol* 1994;14:3166-3175
97. Hirose T, Smith RJ, Jetten AM: ROR gamma: the third member of ROR/RZR orphan receptor subfamily that is highly expressed in skeletal muscle. *Biochem Biophys Res Commun* 1994;205:1976-1983
98. Brickell PM, Tickle C: Morphogens in chick limb development. *Bio Essays* 1995;11:145-149

99. Boylan JF, Lufkin T, Achkar CC, Taneja R, Chambon P, Gudas LJ: Targeted disruption of retinoic acid receptor alpha (RAR alpha) and RAR gamma results in receptor-specific alterations in retinoic acid-mediated differentiation and retinoic acid metabolism. *Mol Cell Biol* 1995;15:843-851
100. Kastner P, Grondona JM, Mark M, Gansmuller A, LeMeur M, Decimo D, Vonesch JL, Dolle P, Chambon P: Genetic analysis of RXR alpha developmental function: convergence of RXR and RAR signaling pathways in heart and eye morphogenesis. *Cell* 1994;78:987-1003
101. Sucov HM, Dyson E, Gumeringer CL, Price J, Chien KR, Evans RM: RXRα mutant mice establish a genetic basis for vitamin A signaling in heart morphogenesis. *Genes Dev* 1994;8:1007-1018
102. Mangelsdorf DJ, Borgmeyer U, Heyman RA, Zhou JY, Ong ES, Oro AE, Kakizuka A, Evans RM: Characterization of three RXR genes that mediate the action of 9-cis retinoic acid. *Genes Dev* 1992;6:329-344
103. Dyson E, Sucov HM, Kubalak SW, Schmid Schonbein GW, DeLano FA, Evans RM, Ross J, Jr., Chien KR: Atrial-like phenotype is associated with embryonic ventricular failure in retinoid X receptor alpha -/- mice. *Proc Natl Acad Sci U S A* 1995;92:7386-7390
104. Morkin E: Regulation of myosin heavy chain genes in the heart. *Circulation* 1993;87:1451-1460
105. Lompre AM, Nadal-Ginard B, Mahdavi V: Expression of cardiac ventricular α- and β myosin heavy chain genes is developmentally and hormonally regulated. *J Biol Chem* 1984;259:6437-6446
106. Milano CA, Allen LF, Rockman HA, Dolber PC, McMinn TR, Chien KR, Johnson TD, Bond RA, Lefkowitz RJ: Enhanced myocardial function in transgenic mice overexpressing the beta 2-adrenergic receptor. *Science* 1994;264:582-586
107. Pollock R, Treisman R: Human SRF-related proteins: DNA-binding properties and potential regulatory targets. *Genes Dev* 1991;5:2327-2341
108. Lilly B, Zhao B, Ranganayakulu G, Paterson BM, Schulz RA, Olson EN: Requirement of MADS domain transcription factor D-MEF2 for muscle formation in Drosophila. *Science* 1995;267:688-693
109. Bour BA, O'Brien MA, Lockwood WL, Goldstein ES, Bodmer R, Taghert PH, Abmayr SM, Nguyen HT: Drosophila MEF2, a transcription factor that is essential for myogenesis. *Genes Dev* 1995;9:730-741
110. Kaushal S, Schneider JW, Nadal Ginard B, Mahdavi V: Activation of the myogenic lineage by MEF2A, a factor that induces and cooperates with MyoD. *Science* 1994;266:1236-1240
111. Edmondson DG, Lyons GE, Martin JF, Olson EN: Mef2 gene expression marks the cardiac and skeletal muscle lineages during mouse embryogenesis. *Development* 1994;120:1251-1263

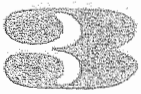
112. Breitbart RE, Liang CS, Smoot LB, Laheru DA, Mahdavi V, Nadal Ginard B: A fourth human MEF2 transcription factor, hMEF2D, is an early marker of the myogenic lineage. *Development* 1993;118:1095-1106
113. Leifer D, Krainc D, Yu YT, McDermott J, Breitbart RE, Heng J, Neve RL, Kosofsky B, Nadal Ginard B, Lipton SA: MEF2C, a MADS/MEF2-family transcription factor expressed in a laminar distribution in cerebral cortex. *Proc Natl Acad Sci U S A* 1993;90:1546-1550
114. Chambers AE, Kotecha S, Towers N, Mohun TJ: Muscle-specific expression of SRF-related genes in the early embryo of Xenopus laevis. *EMBO J* 1992;11:4981-4991
115. Martin JF, Miano JM, Hustad CM, Copeland NG, Jenkins NA, Olson EN: A Mef2 gene that generates a muscle-specific isoform via alternative mRNA splicing. *Mol Cell Biol* 1994;14:1647-1656
116. McDermott JC, Cardoso MC, Yu YT, Andres V, Leifer D, Krainc D, Lipton SA, Nadal Ginard B: hMEF2C gene encodes skeletal muscle- and brain-specific transcription factors. *Mol Cell Biol* 1993;13:2564-2577
117. Martin JF, Schwarz JJ, Olson EN: Myocyte enhancer factor (MEF) 2C: a tissue-restricted member of the MEF-2 family of transcription factors. *Proc Natl Acad Sci U S A* 1993;90:5282-5286
118. Zhu H, Nguyen VT, Brown AB, Pourhosseini A, Garcia AV, van Bilsen M, Chien KR: A novel, tissue-restricted zinc finger protein (HF-1b) binds to the cardiac regulatory element (HF-1b/MEF-2) in the rat myosin light-chain 2 gene. *Mol Cell Biol* 1993;13:4432-4444
119. Kelley C, Blumberg H, Zon LI, Evans T: GATA-4 is a novel transcription factor expressed in endocardium of the developing heart. *Development* 1993;118:817-827
120. Molkentin JD, Kalvakolanu DV, Markham BE: Transcription factor GATA-4 regulates cardiac muscle-specific expression of the alpha-myosin heavy-chain gene. *Mol Cell Biol* 1994;14:4947-4957
121. Arceci RJ, King AA, Simon MC, Orkin SH, Wilson DB: Mouse GATA-4: a retinoic acid-inducible GATA-binding transcription factor expressed in endodermally derived tissues and heart. *Mol Cell Biol* 1993;13:2235-2246
122. Heikinheimo M, Scandrett JM, Wilson DB: Localization of transcription factor GATA-4 to regions of the mouse embryo involved in cardiac development. *Dev Biol* 1994;164:361-373
123. Laverriere AC, MacNeill C, Mueller C, Poelmann RE, Burch JB, Evans T: GATA-4/5/6, a subfamily of three transcription factors transcribed in developing heart and gut. *J Biol Chem* 1994;269:23177-23184
124. Ip HS, Wilson DB, Heikinheimo M, Tang Z, Ting C, Simon M, Leiden J, Parmacek MS: The GATA-4 transcription factor transactivates the cardiac muscle-specific troponin C promoter-enhancer in nonmuscle cells. *Mol Cell Biol* 1994;14:7517-7526
125. Thuerauf DJ, Hanford DS, Glembotski CC: Regulation of rat brain natriuretic peptide transcription. A potential role for GATA-related transcription factors in myocardial cell gene expression. *J Biol Chem* 1994;269:17772-17775

126. Mar JH, Ordahl CP: M-CAT binding factor, a novel trans-acting factor governing muscle-specific transcription. *Mol Cell Biol* 1990;10:4271-4283
127. Farrance IK, Mar JH, Ordahl CP: M-CAT binding factor is related to the SV40 enhancer binding factor, TEF-1. *J Biol Chem* 1992;267:17234-17240
128. Shimizu N, Smith G, Izumo S: Both a ubiquitous factor mTEF-1 and a distinct muscle-specific factor bind to the M-CAT motif of the myosin heavy chain beta gene. *Nucleic Acids Res* 1993;21:4103-4110
129. Stewart AF, Larkin SB, Farrance IK, Mar JH, Hall DE, Ordahl CP: Muscle-enriched TEF-1 isoforms bind M-CAT elements from muscle-specific promoters and differentially activate transcription. *J Biol Chem* 1994;269:3147-3150
130. Karns LR, Kariya K, Simpson PC: M-CAT, CArG, and Sp1 elements are required for alpha 1-adrenergic induction of the skeletal alpha-actin promoter during cardiac myocyte hypertrophy. Transcriptional enhancer factor-1 and protein kinase C as conserved transducers of the fetal program in cardiac growth. *J Biol Chem* 1995;270:410-417
131. Kariya K, Karns LR, Simpson PC: 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* 1994;269:3775-3782
132. MacLellan WR, Lee TC, Schwartz RJ, Schneider MD: Transforming growth factor-beta response elements of the skeletal alpha-actin gene. Combinatorial action of serum response factor, YY1, and the SV40 enhancer-binding protein, TEF-1. *J Biol Chem* 1994;269:16754-16760
133. Chen L, Krause M, Sepanski M, Fire A: The Caenorhabditis elegans MYOD homologue HLH-1 is essential for proper muscle function and complete morphogenesis. *Development* 1994;120:1631-1641
134. Knotts S, Rindt H, Neumann J, Robbins J: In vivo regulation of the mouse beta myosin heavy chain gene. *J Biol Chem* 1994;269:31275-31282
135. Lee TC, Chow KL, Fang P, Schwartz RJ: Activation of skeletal alpha-actin gene transcription: the cooperative formation of serum response factor-binding complexes over positive cis-acting promoter serum response elements displaces a negative-acting nuclear factor enriched in replicating myoblasts and nonmyogenic cells. *Mol Cell Biol* 1991;11:5090-5100
136. Chow KL, Schwartz RJ: A combination of closely associated positive and negative cis-acting promoter elements regulates transcription of the skeletal alpha-actin gene. *Mol Cell Biol* 1990;10:528-538
137. Min BH, Foster DN, Strauch AR: The 5'-flanking region of the mouse vascular smooth muscle alpha-actin gene contains evolutionarily conserved sequence motifs within a functional promoter. *J Biol Chem* 1990;265:16667-16675

138. Kurabayashi M, Komuro I, Shibasaki Y, Tsuchimochi H, Takaku F, Yazaki Y: Functional identification of the transcriptional regulatory elements within the promoter region of the human ventricular myosin alkali light chain gene. *J Biol Chem* 1990;265:19271-19278
139. Klamut HJ, Gangopadhyay SB, Worton RG, Ray PN: Molecular and functional analysis of the muscle-specific promoter region of the Duchenne muscular dystrophy gene. *Mol Cell Biol* 1990;10:193-205
140. Mohun TJ, Chambers AE, Towers N, Taylor MV: Expression of genes encoding the transcription factor SRF during early development of Xenopus laevis: identification of a CArG box-binding activity as SRF. *EMBO J* 1991;10:933-940
141. Argentin S, Ardati A, Tremblay S, Lihrmann I, Robitaille L, Drouin J, Nemer M: Developmental stage-specific regulation of atrial natriuretic factor gene transcription in cardiac cells. *Mol Cell Biol* 1994;14:777-790
142. Bassel Duby R, Hernandez MD, Gonzalez MA, Krueger JK, Williams RS: A 40-kilodalton protein binds specifically to an upstream sequence element essential for muscle-specific transcription of the human myoglobin promoter. *Mol Cell Biol* 1992;12:5024-5032
143. Bassel Duby R, Hernandez MD, Yang Q, Rochelle JM, Seldin MF, Williams RS: Myocyte nuclear factor, a novel winged-helix transcription factor under both develop-mental and neural regulation in striated myocytes. *Mol Cell Biol* 1994;14:4596-4605
144. Wang G, Yeh HI, Lin JJ: Characterization of cis-regulating elements and trans-activating factors of the rat cardiac troponin T gene. *J Biol Chem* 1994;269:30595-30603.
145. Vincent CK, Gualberto A, Patel CV, Walsh K: Different regulatory sequences control creatine kinase-M gene expression in directly injected skeletal and cardiac muscle. *Mol Cell Biol* 1993;13:1264-1272
146. Zou Y, Chien KR: EFIA/YB-1 is a component of cardiac HF-1A binding activity and positively regulates transcription of the myosin light-chain 2v gene. *Mol Cell Biol* 1995;15:2972-2982
147. Molkentin JD, Markham BE: Myocyte-specific enhancer-binding factor (MEF-2) regulates alpha-cardiac myosin heavy chain gene expression in vitro and in vivo. *J Biol Chem* 1993;268:19512-19520
148. Molkentin JD, Brogan RS, Jobe SM, Markham BE: Expression of the alpha-myosin heavy chain gene in the heart is regulated in part by an E-box-dependent mechanism. *J Biol Chem* 1993;268:2602-2609
149. Kelly R, Alonso S, Tajbakhsh S, Cossu G, Buckingham M: Myosin light chain 3F regulatory sequences confer regionalized cardiac and skeletal muscle expression in transgenic mice. *J Cell Biol* 1995;129:383-396
150. Verrijzer PC, Yokomori K, Chen J, Tjian R: *Drosophila* $TAF_{II}150$: similarity to yeast gene TSM-1 and specific binding to core promoter DNA. *Science* 1994;264:933-941
151. Buckingham M: Molecular biology of muscle development. *Cell* 1994;78:15-21.

152. Thisse B, Stoetzel C, Gorostiza-Thisse C, Perrin-Schmitt F: Sequence of the twist gene and nuclear localization of its protein in endo-mesodermal cells of early *Drosophila* embryos. *EMBO J* 1988;7:2175-2183.
153. Azpiazu N, Frasch M: *Tinman* and *bagpipe:* two homeobox genes that determine cell fates in the dorsal mesoderm of *drosophila. Genes Dev* 1993;7:1325-1340
154. Donoviel DB, Shield MA, Buskin JN, Haugen HS, Clegg CH, Hauschka SD: Analysis of muscle creatine kinase gene regulatory elements in skeletal and cardiac muscles of transgenic mice. *Mol Cell Biol* 1996;16:1649-1658.
155. Yu VC, Delsert C, Andersen B, Holloway JM, Devary OV, Naar AM, Kim SY, Boutin JM, Glass CK, Rosenfeld MG: RXR beta: a coregulator that enhances binding of retinoic acid, thyroid hormone, and vitamin D receptors to their cognate response elements. *Cell* 1991;67:1251-1266
156. Ogura T, Evans RM: Evidence for two distinct retinoic acid response pathways for HOXB1 gene regulation. *Proc Natl Acad Sci U S A* 1995;92:392-396
157. Studer M, Popperl H, Marshall H, Kuroiwa A, Krumlauf R: Role of a conserved retinoic acid reponse element in rhombomere restriction of Hoxb-1. *Science* 1994;265:1728-1732
158. Yu W, Dahl G, Werner R: The connexin 43 gene is responsive to oestrogen. *Proc R Soc Lond B Biol Sci* 1994;255:125-132
159. Kadonaga JT, Courey AJ, Ladika J, Tijan R: Distinct regions of Sp1 modulate DNA binding and transcriptional activation. *Science* 1988;242:1566-1569
160. Gualberto A, LePage D, Pons G, Mader SL, Park K, Atchison ML, Walsh K: Functional antagonism between YY1 and the serum response factor. *Mol Cell Biol* 1992;12:4209-4214
161. Minty A, Kedes L: Upstream regions of the human cardiac actin gene that modulate its transcription in muscle cells: presence of an evolutionary conserved repeated motif. *Mol Cell Biol* 1986;6:2125-2136
162. Johnson JE, Wold JB, Hauschka DD: Muscle creatine kinase sequence elements regulating skeletal and cardiac muscle expression in transgenic mice. *Mol Cell Biol* 1989;9:3393-3399.
163. Chambers AE, Logan M, Kotecha S, Towers N, Sparrow D, Mohun TJ: The RSRF/MEF2 protein SL1 regulates cardiac muscle specific transcription of a myosin light-chain gene in *Xenopus* embryos. *Genes Dev* 1994;8:1324-1334

Chapter



Hunting down nucleic acid binding factors in the cardiovascular system

Pieter A. Doevendans, Pilar Ruiz-Lozano, Marc van Bilsen
Review for Cardiovascular Research: in press

Abstract

Transcription regulation of genes active in the cardiovascular system is a complex process, involving DNA and RNA binding proteins. Nucleic acid binding proteins bind to the regulatory DNA and interact with other proteins, including RNA polymerase to initiate and control the level of transcription. The RNA binding proteins have a function in spliceosome formation and in stabilizing mRNA. In this review the currently available molecular approaches to analyze regulatory DNA in relation to DNA binding proteins are discussed. Similar techniques that have been developed for RNA binding protein studies are included. In addition to an explanation of the various methods, examples are provided from DNA-protein interactions on genes active in the cardiovascular system. Furthermore strategies for identification and characterization of new nucleic acid binding proteins active in cardiomyocytes or vascular cell types are presented.

Introduction

The basic principles of modern molecular biology have been introduced in Cardiovascular Research by Swynghedauw and co-workers in two instructive reviews.[1,2] These papers were complemented with overviews by van Bilsen & Chien and by Robbins on the mechanisms of cardiac transcription regulation during cardiac development and hypertrophy.[3,4] Key players both on the DNA and protein level in muscle formation and cardiomyocyte differentiation were addressed, illustrating the complexity of transcription regulation. In this paper we will make use of terminology explained in the previous reviews, while discussing the currently available molecular techniques to identify and characterize novel genes involved in the process of cardiovascular gene transcription and post transcriptional regulation.

The general concept of gene structure distinguishes regulatory DNA and transcribed DNA regions. The transcribed DNA can be separated in exons (coding DNA) and introns. The regulatory DNA is composed of the 5' flanking region containing the so called essential or core promoter and enhancers/silencers positioned before the first exon. The regulatory DNA can also involve the first intron and the 3' flanking region located behind the coding region.[2,5,6] To initiate transcription the localization and orientation of the essential promoter is fixed. In contrast the position of enhancer/silencer sequences is less critical which is illustrated experimentally by the fact that these sequences can be moved throughout the gene and their orientation can be reversed, without loosing function. The rate of transcription of a particular gene is determined by a complex of transcription (*trans*-acting) factors binding to short DNA sites (4-12 bases) with a conserved core nucleotide sequence (*cis*-elements) located in promoter and enhancer/silencer regions. An enhancer/silencer region can encompass one or several *cis* elements. The muscle specific enhancer of the creatine kinase gene, active in both skeletal and cardiac muscle, contains several DNA binding core sequences, each interacting with specific transcription factors.[5] Both DNA-protein and protein-protein interaction of various transcription factors, play an important role in transcription regulation.

Processing of the primary transcript is a second regulating mechanism for gene expression. Splicing of the primary transcript and the formation of mRNA requires the cooperation of a set of nucleic acid binding proteins (NABP) forming spliceosomes. The formation of a spliceosome requires RNA-protein interaction, and conserved RNA recognition sites at the intron-exon boundaries. Two groups of protein containing complexes are involved in spliceosome formation, i.e. the conserved small nuclear ribonucleoprotein and the heterogeneous ribonucleoprotein complexes. Both play a role in the preservation of single stranded RNA and the active process of splicing. The process of RNA splicing involves capping of the 5' end and the addition of a poly-adenine nucleotide tail at the 3' end. The

final mRNA product is exported from the nucleus and forms the template for ribosomal protein synthesis.

The study of the mechanisms that regulate cardiac specific expression has been difficult due to the lack of stable cell lines expressing differentiated cardiac markers. However the study of the expression of contractile proteins has long been a paradigm in cardiac gene regulation. In the developing embryo several cardiac proteins change from fetal to adult isoforms.[7,8] The transition from the beating heart tube to the mature cardiac system involves the expression of different myofibrillar, membrane and cytosolic protein isoforms. This developmental switch from fetal to adult is mediated both by differential splicing of single primary transcripts and transcriptional induction of the adult isoforms.[9-11]

In the past decade different methods have been developed to study nucleic acid-protein interaction, both *in vivo* and *in vitro*. Molecular techniques enable us to study the exact DNA sequence required for transcription factor binding and the importance of single base mutations in a DNA *cis* element.[12] Similar experiments can be performed to study RNA-protein interaction in spliceosomes.[13] The NABP are classified according to the protein domain directly interacting with either DNA or RNA. NABP interact either with single stranded (RNA and DNA), or double stranded nucleic acid molecules (DNA), or both.[14-17] In this chapter we will outline different methods that can be used to study nucleic acid-protein interaction and illustrative examples will be provided. As the starting point it is assumed that a genomic clone of interest is available, and that the main aim of the investigator is the characterization of the regulatory, 5' flanking region of the gene.[2] The methodology that is used to identify new NABP and to unravel mechanisms involved in transcription regulation of the selected gene are discussed. When available, genes active in the cardiovascular system will be used to clarify the various methods.

Identification of the transcription start site

The transcription start site marks the physical separation between the essential promoter and the transcribed gene. The transcription start site provides the first base of the RNA molecule. Although similar patterns of transcription initiation are found in various genes, there are no rigid biological laws. Many genes have a conserved *cis* element, with the consensus sequence TATA(A/T)A(A/T) (TATA-box). This is the binding site for the TATA-box binding protein, part of the transcription initiation protein complex, needed for RNA polymerase II binding. Transcription starts approximately 30 bp downstream (3') of the TATA box.[2] Other genes have a different conserved element surrounding the transcription start site. This element was named Initiator (Inr) and the consensus sequence is CTCANTCT.[12] Inr sequences were initially described in housekeeping genes, but

recently it has been shown that highly regulated tissue specific genes also contain Inr sequences in the absence of a TATA box. For example the murine atrial isoform of the regulatory myosin light chain (MLC-2a) gene is completely silenced after partial or complete deletion of the Inr element.[18] The identification of the transcription initiation site of a cloned gene is an important step in the analysis of a new gene. Several methods can be applied to determine the start site.

RNase protection: For this approach the complementary RNA strand is synthesized from a complementary DNA (cDNA) containing plasmid and allowed to interact with the mRNA molecules. The ends of the partially hybridized molecule that remains single stranded can be digested by RNase treatment and the size of the protected (double stranded) fragment determined by gel electrophoresis (figure 1A). This method can also be applied to determine intron-exon boundaries and to establish the localization, size, and number of distinctive exons.

Primer extension: This method is based on hybridization of an oligonucleotide to the mRNA and subsequent filling of the proximal mRNA tail (5'). The oligonucleotide used is designed in the antisense orientation and specifically recognizes the mRNA of interest. This oligonucleotide is used as a primer to synthesize a complementary DNA strand end labeled and by adding the enzyme reverse transcriptase, a RNA-dependent DNA polymerase. The size of the extended fragment can be analyzed by running the primer extension reaction product in parallel to a sequencing reaction of the genomic clone. The base that matches the primer extended product indicates the start site (a thymidine residue in figure 1B). In some genes several extended bands can be detected, indicating the presence of various transcription start sites and variation in the mRNA length. For instance primer extension studies of the Shaker-like potassium channel gene (Kv1.5) revealed three products, differing 118 and 164 bp in size.[7] This method has been modified and optimized and is now available in a specially designed kit. This 5' rapid amplification of cDNA ends kit requires, in addition to primer extension, the modification of the extended product by a poly-C tail, recognized by a special anchor primer. The first strand reverse transcriptase product can be amplified by polymerase chain reaction (PCR) and the product can be cloned and sequenced (figure 2).

If the cDNA clone is available and the sequence of the exons is known, it can be used to design oligonucleotides for PCR and Southern blotting. Both tech-niques can be used to estimate the localization and size of introns. For an exact determination of the intron-exon boundaries double strand (both DNA strands) sequencing is required. Introns have conserved 5' (GU) and 3' (AG) sequences and often a pyrimidine (CU) rich region close to the 3' end,[13] however the limited length of the consensus sequences does not unambiguously indicate the

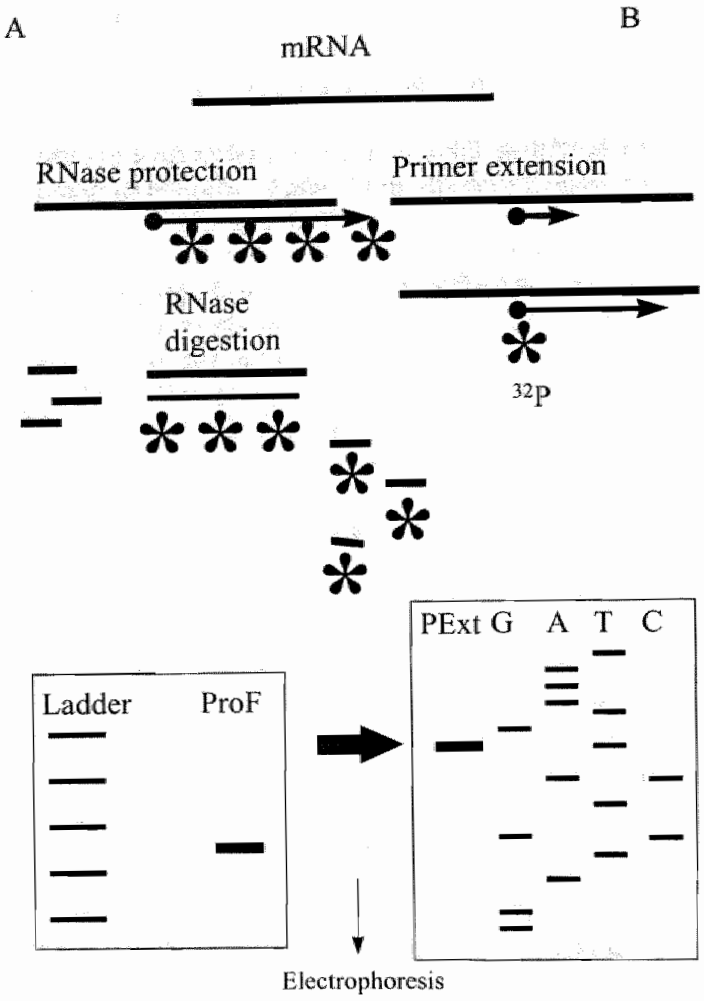


Figure 1. *Primer extension and RNase protection:* Schematic representation of experiments to identify the transcription start site.On the left side (A) the principle of the RNA protection method is shown schematically. For the experiment tissue derived mRNA, and a plasmid containing the genomic clone of interest are needed. From the genomic clone a RNA probe can be synthesized by using RNA polymerase (SP6 or T7), which is homogeneously labeled by adding an isotope labeled nucleotide. The synthesized RNA molecule should be long enough to encompass the expected transcription start sites. Following hybridization of probe with mRNA, the molecule is digested by RNase, leaving the double strand molecule intact. The size of the protected RNA-RNA hybrid regions, determined by electrophoresis in the presence of a DNA ladder, indicates the size of the mRNA and thus the transcription start site. ProF: protected fragment. The right panel (B) outlines the method of primer extension. An end labeled antisense primer is hybridized to a specific mRNA molecule and in the presence of Klenow enzyme fragment the gap is filled in the direction of the 5' flanking region. The extension reaction product is run simultaneously with a regular sequencing reaction using the same primer on genomic DNA. In this example transcription starts at a thymidine nucleotide, indicated by black arrow. (*) indicates radioactive label. PExt: primer extension product.

presence of introns. Up to now sequence comparison between genomic DNA and cDNA, is the only way to solve this issue.

Identification of cis acting elements

Sequence comparison: After determination of the transcription start site, the putative regulatory region can be sequenced using both DNA strands. Sequence comparison of selected genes from different species usually reveals large variations in the regulatory 5' flanking region, whereas the coding region is often highly conserved. However in the 5' flanking sequence some short DNA sequences appear to be conserved regarding both orientation and position relative to the transcription start site. Evolutionary conservation of such DNA sequences points to their functional importance. This type of analysis has frequently been used to identify DNA regions of potential importance in transcriptional regulation. This method of sequence comparison (table 1) was used successfully to identify important sites in the ventricular myosin light chain 2 promoter (MLC-2v).[19,20] The sequence of chick MLC-2v was compared to the sequence of the rat clone and conserved sites were recognized and named HF1, HF2 and HF3. These DNA elements were further explored and appeared to be important for ventricular cardiac myocyte specific expression. The regulatory sequences of different genes within one species can also be used for sequence comparison to identify potentially interesting cis elements. Comparison of rat Kv1.5 regulatory sequence with the somatostatin regulatory DNA revealed the presence of a consensus cAMP responsive element (CRE, sequence: TGACGTCA).[21] If no related sequences are available for comparison, specialized computer software has been developed to scan a sequence for conserved *cis* elements.

Transient transfections assays: The functional relevance of a regulatory DNA sequence can be determined by using reporter genes in transfections of mammalian cells (table 1). The common feature of reporter genes is the availability of a reliable and sensitive method to measure the protein product of the gene. By subcloning (parts of) a promoter sequence upstream of the reporter gene, the expression of the reporter gene will be under control of this promoter fragment and, hence, provide a measure of the activity of the promoter fragment. Reporter genes commonly used are chloramphenicol acetyl transferase (CAT)[22] and luciferase[19,23] which can be assayed quantitatively. The β-galactosidase reporter gene can be used for both qualitative and quantitative assays.[19,24] The level of expression can be compared to well characterized viral promoters like the human cytomegalovirus (CMV) and the Rous-Sarcoma virus (RSV), and thus provide a relative measure. Not seldomly transfections of cultured cells show considerable variability in consecutive experiments. Therefore it is common-use to co-transfect cells with a luciferase vector, containing the promoter fragment of interest,

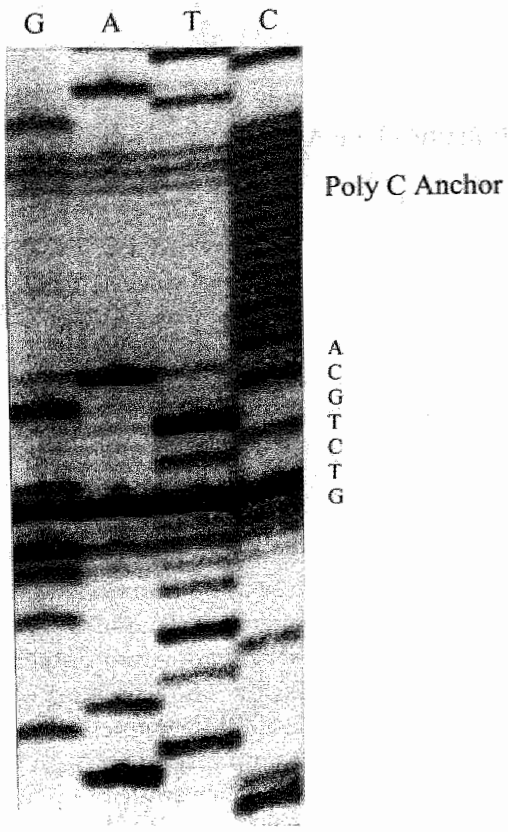


Figure 2. RACE rt-PCR: Picture of a sequencing reaction of polymerase chain reaction product generated by the rapid amplification of cDNA ends method (5' RACE, Gibco BRL, Gaithersburg MD). This method to identify the transcription start site was used on a genomic murine atrial myosin light chain 2 clone. The method briefly: for first strand cDNA synthesis a gene specific primer is used approximately 300 bp from the cDNA end (ATGAGAAGCTGCTTGAACTC). An anchor sequence is then added to the 3' end of the first strand DNA (complementary to the 5' end of the mRNA) using homopolymeric tailing with dC. The dC tail is recognized by a specially designed anchor primer and this primer is used in a PCR employing a second gene specific (nested) antisense primer (GAATAGGTCTCCTTCAGGTC) and the *Taq* DNA polymerase. The PCR product was cloned into the TA cloning vector (Stratagene, San Diego, CA), and sequenced according to the method of Sanger.[93] The picture shows the first bases of the cDNA corresponding to the 5' end of the mRNA (TCTGCAC) preceded by the poly C anchor.

Table 1. Strategies to study *cis*-acting nucleic acid elements and *trans*-acting nucleic acid binding proteins.

1) *Identification of* cis *elements*
   A. Sequence comparison[19,89]
   B. Functional assays: Transfection of the complete or partial regulatory sequence linked to a reporter gene.[19,30] *In vitro* studies with precursor mRNA and spliceosome related extracts.[36]
2) *Characterization of* cis *elements*
   A: Binding assays: DNase I footprinting[37], Gel mobility shift assay[39], Methylation interference *in vitro* and *in vivo*[39,40]
   B. Functional assays: Transfection of wild type versus mutated *cis* element linked to a reporter gene.[19] Transgenic mice harboring wild type or mutated regulatory sequences linked to a reporter gene.[20]
3) *Cloning of nucleic acid binding proteins*
   A. Degenerate primer polymerase chain reaction[50]
   B. Screening of cDNA library[58]
   C. Protein purification, sequencing and conversion into DNA sequence[60]
   D. Southwestern screening[61]
   E. Filter mediated NABP domain-antibody interaction
4) *Characterization of nucleic acid binding proteins*
   A. Sequence comparison [64]
   B. Protein structure analysis[71]
   C. Expression pattern: Northern blotting, in situ hybridization[70]
5) *Functional analysis of nucleic acid binding proteins*
   A. Functional assays: Co-transfection of DNA binding protein sequence with wild-type regulatory sequence linked to reporter.[62]
   B. Gene targeting of the nucleic acid binding factor.[90]

together with a β-galactosidase vector driven by the CMV promoter, that can be used to correct for differences in transfection efficiency.[25] The green fluorescent protein originally detected in jellyfish, *Aequorea victoria*, can be used as an in vivo reporter gene in mammalian cells.[26-28] This reporter gene can be used to identify living transfected cells, and the tools for quantification of the level of expression in vital cells are being developed. Mutations in the green fluorescent protein have been made and appeared to generate a stronger fluorescent signal.[26,28] In addition through further manipulation of the gene, other color producing fluorescent proteins have been developed.[29] These reporter genes are commercially available in plasmids with convenient multiple cloning sites 5' of the reporter cDNA.

After subcloning of the regulatory sequence upstream of the reporter gene, large quantities of recombinant plasmid DNA can be harvested from bacterial cultures. The plasmid DNA is then purified and transfected into the desired cell type. This approach has also been used to study regulatory sequences of ion

channel genes. Constructs of the Kv1.5 regulatory DNA driving the CAT reporter gene, were used to study the change in transcription level after cAMP treatment in the presence of a conserved CRE and mutated CRE site in skeletal myocytes.[7] Functionally important *cis* elements can be identified by studying deletion fragments of the regulatory sequence. For these experiments the deletion fragments are cloned upstream of a reporter gene and variations in the level of reporter gene activity are measured and indicative for either repressive or stimulating *cis* elements in the deleted fragments. In general as the deleted fragments are large (several hundred basepairs in length) this method is used primarily to define DNA regions containing potentially interesting regulatory sites.[30] The choice of reporter gene and cell type will depend on the gene under investigation. Neonatal rat cardiomyocytes have been used extensively as a model system to study transcription regulation in the cardiac context. The cardiomyocytes can be transiently transfected by direct DNA injection or by using either calcium phosphate precipitation, liposomes, or adenovirus mediated transfection.[31,32] The transfection method will have to be adjusted to the strength of the regulatory sequence in inducing reporter gene expression. The calcium phosphate precipitation method is relatively easy, but the efficiency of transfection is low (<5%). In contrast by using replication deficient adenovirus near 100% transfection can be obtained, but the generation of adenoviral constructs is time consuming and viral DNA may influence reporter gene expression levels. A distinct advantage of adenovirus as a vehicle to introduce foreign DNA is that it also allows transfection of adult myocytes.[33]

In vitro spliceosome assays: The formation of mRNA, capable of crossing the nuclear membrane, is dependent on the correct function of RNA-protein complexes. Proteins recognize the boundaries between introns and exons (localization of splice sites), cut the precursor RNA molecule, and paste the individual exons in a fixed order (splicing). The function of the introns remains obscure for most genes. In some cases the introns were shown to contain tissue specific enhancers.[5] Whether this enhancer function is mediated by DNA-protein interaction prior to, or by RNA-protein interactions following transcription is unclear. A defect in the splicing process forms the molecular bases for the expression of a defective Ca^{2+} channel β-subunit, resulting in ataxia and seizures in the lethargic mice.[34] During the aberrant splicing events exon III of the precursor RNA of the β-subunit is lost, and not included in the final mRNA molecule. The resulting mutated splice variant is characterized by a frame shift, which leads in this particular example to stop codon formation and results in a truncated protein. In the human ether a-go-go related gene (Herg), a similar splice donor mutation (G to C mutation at the splice site of intron III)[35] leads to protein changes in the Herg ion channel and modification of the I_{Kr} current that can cause the long QT syndrome.

By mixing precursor RNA with nuclear extracts or with purified protein mixtures containing the spliceosome forming proteins the formation of mRNA can be assessed by studying the length and structure of the various RNA molecules, using primer extension (figure 1). Through this method the functional significance of mutations surrounding the splice sites can be determined. This method has been used extensively to show the role of competing splice factors in gender determination in *Drosophila*.[36] A similar approach was used by Mullen et al.[13] to study the mechanism of exon selection of the α- Tropomyosin gene. The mRNA formed in smooth muscle cells includes exon 2 and skips exon 3, where other cell types skip exon 2 and paste exon 1 to 3. The investigators showed that these splicing events can be reproduced *in vitro*, by using Hela cell nuclear extracts. In this study an important role for the polypyrimidine tract adjacent to the splice sites in exon selection was found.

Characterization of cis elements

For the characterization of *cis* elements various methods can be applied, some of which are complementary. The first approach is using nucleic acid-protein binding assays to determine the exact DNA sequence that forms a nucleo-protein complex, and the strength of the interaction *in vitro*. The techniques of DNA footprinting, gel mobility shift assay and methylation interference can be applied to address these questions. The effect of mutations in a conserved element can be compared with the normal (wildtype) sequence in transient transfections (*in vitro)* and through the generation of transgenic mouse lines (*in vivo*, table 2)

DNA footprinting: Schmitz et al.[37] described a technique to show which bases were protected against endonuclease enzyme DNase I digestion. Nuclear extract of interest is incubated with a regulatory DNA fragment, labeled at one end, in a binding reaction buffer. After DNA-protein binding the DNA can be digested briefly, and thus incompletely, by DNase I. Simultaneously labeled DNA free of protein is digested as a control. The unbound free DNA will yield a characteristic pattern of bands on electrophoresis on a standard sequencing gel. The NABP generate a so-called footprint when the DNase I digested DNA bands are analyzed, as the bands resulting from free DNA digestion will be missing at those DNA sequences, that were protected from DNase I digestion through their occupancy by proteins (figure 3, 4). By DNase I footprinting DNA fragments up to 250 bp can be studied, thus resulting in a broad overview of site occupation.

Gel mobility shift assay: The DNA nucleotides involved in protein interaction and the strength of binding can be determined by the gel mobility shift assay (GMSA), also referred to as electrophoretic mobility shift assay. This technique is based on electrophoresis, where a free unbound labeled DNA probe will move through a gel faster compared to a protein bound DNA molecule. The delay in

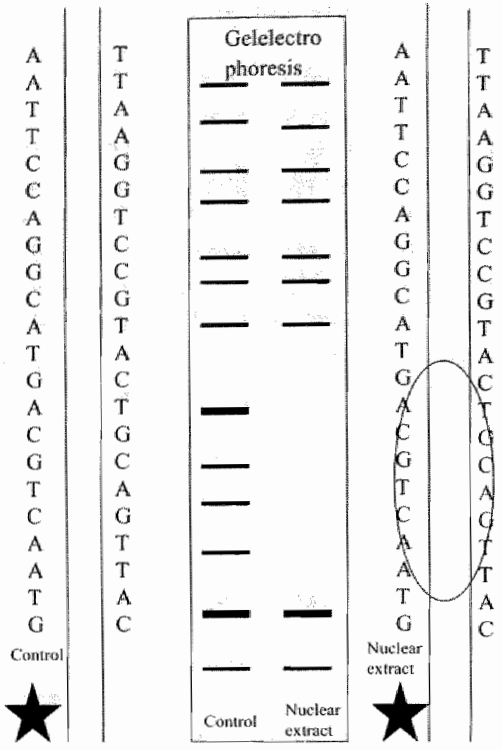


Figure 3. DNase I footprinting: Scheme showing the basics of DNA footprinting. The transparent oval indicates a nucleic acid binding factor interacting with the region TCACGTCA (*cis* element). One DNA strand is end labeled with T4 poly-nucleotide kinase and γ^{32}P-ATP, and the DNA-protein complex is subjected to brief DNase I digestion. The resulting fragments are analyzed by gel electro-phoresis. The control experiment is performed in the absence of nuclear extract.

electrophoretic mobility helps to identify the presence of factors binding to the DNA site of interest. The source of proteins can be either a crude total cell extract, an isolated nuclear extract, a mammalian protein produced by bacteria, or a protein produced through in vitro transcription and translation. To reduce the background in the DNA-protein interaction, unlabeled nonspecific competitor is often added (figure 5). The DNA-protein interaction is dependent on the quality of the protein extract. As the proteins are sensitive to heat, oxidation, degradation and mechanical shearing, extensive precautions should be taken to preserve protein function. Furthermore the interaction can require specific salt concentrations and co-factors of the binding buffer.[38] To further demonstrate specificity of the interaction, unlabeled DNA identical to the probe can be added. Increasing concentration of the competitor will show a gradual decrease in the binding to the labeled probe. Conversely the presence of increasing concentrations of random DNA fragments, should not influence the binding of the probe. In addition, the influence of mutations in the DNA site can be assessed again by establishing binding affinity and by showing the effects on specific competition (figure 5, 6). If antibodies are available against a NABP, they can be used to show an additional slowing of mobility as the labeled DNA-protein complex is be-

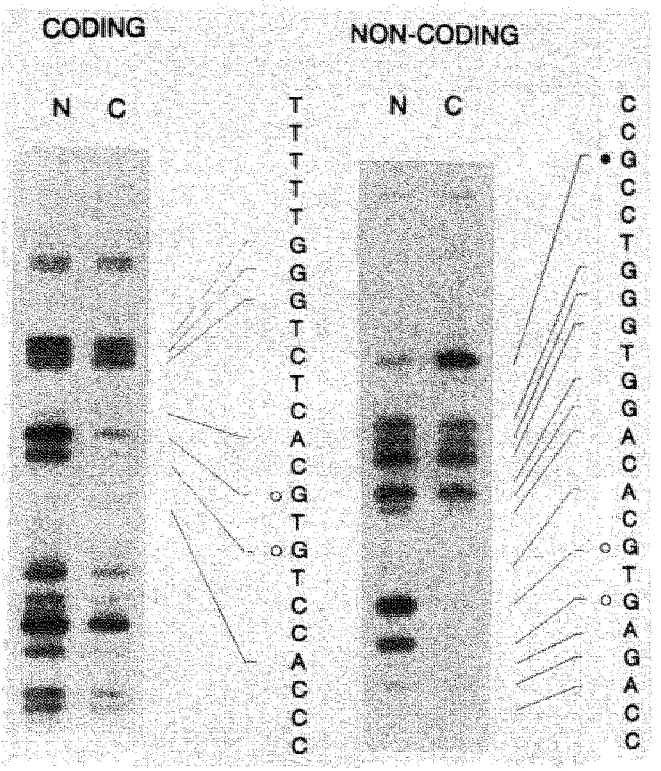


Figure 4. In vivo footprinting: Identification of DNA-protein interaction by guanine or adenine *in vivo* footprinting method based on methylation interference. The genomic DNA is subjected to methylation with dimethyl sulfate *in vivo*. The DNA is then treated with piperidine cleaving DNA directly after a methylated (guanine or adenine) site. In this experiment the DNA fragments were amplified and labeled using PCR. If the protein binding *in vivo* prevented DNA methylation, bands are missing in the *in vivo* methylated DNA compared to the free probe after size fractionation on a denaturing gel (method from Strauss et al. Positions of nucleotides, which are methylation protected following binding to the nuclear factors (°) and hypersensitive site (•). Lanes: N, *in vitro*-methylated protein-free cardiac myocyte DNA; C, *in vivo*-methylated cardiac myocyte DNA. (reproduced from Navankasattusas et al, Mol Cell Biol, 1994: 14: 7336, with permission)

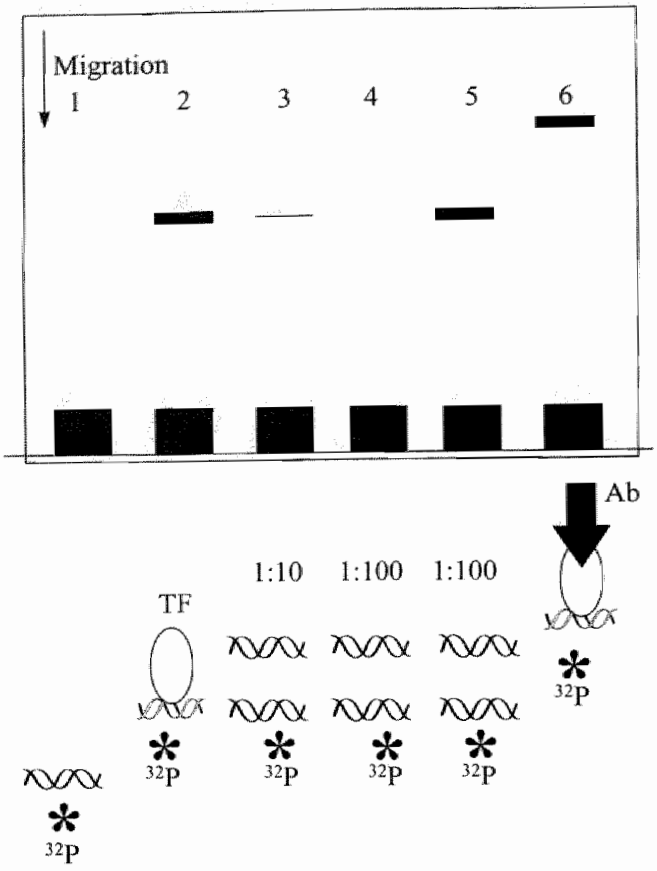


Figure 5. Gel mobility shift assay: Scheme to illustrate the method for gel mobility shift assays. In the first lane the labeled unbound probe is visualized. The second lane shows the shift in mobility of the probe after incubation with nuclear extract. By adding increasing amounts of unlabeled DNA identical to the probe, competition results in diminished intensity of the shifted band (lane 3,4). No competition is seen in lane 5 where non-related or mutated DNA was added. In the last lane specific antibodies are added to the DNA-protein incubation reaction. The antibody recognizes and binds the transcription factor, which is still able to interact with the DNA probe. The resulting DNA-protein complex is larger and therefore electrophoretic mobility of the complex is further retarded.

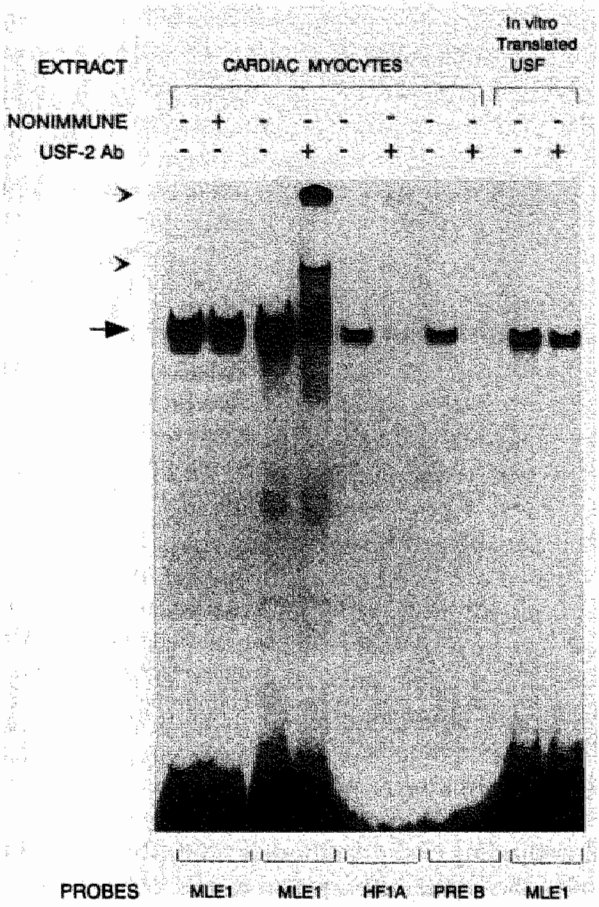


Figure 6. Gel mobility shift assay: Gel mobility shift comparing various *cis* elements (MLE-1, HF1A and PRE B) and protein binding in ventricular cardiomyocyte nuclear extract. The normal band indicating protein-DNA interaction is the lower band. In the presence of the serum containing USF-2 antibodies (USF-2-ab) the band is supershifted (more retarded). In this experiment binding characteristics of USF to different E-box containing sites (MLE-1, HF1A, PRE-B) was studied. Please note that the *in vitro* transcribed USF protein although capable of DNA binding is not recognized by the rabbit serum. (reproduced from Navankasat-tusas et al, Mol Cell Biol, 1994: 14: 7336, with permission)

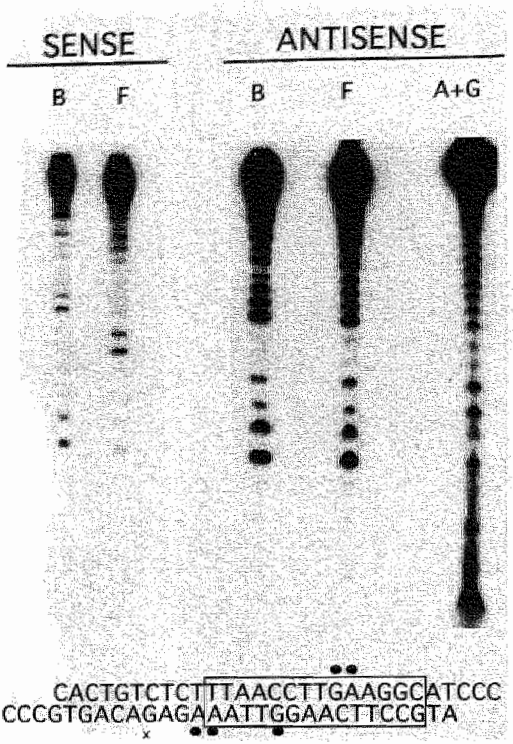


Figure 7. DEPC interference: A NABP was identified (by Southwestern screening), using a conserved HF3 *cis* element as a probe and referred to as cardiac negative regulator (CNR). For this experiment an oligonucleotide containing the HF3 element was end-labeled, either the sense or the antisense strand, and partially modified by DEPC treatment. The DNA-protein complexes formed between the HF3 element and the bacterially expressed GST-CNR fusion protein were resolved and isolated from a preparative 5% acrylamide non-denaturing gel by electroelution. After removal of the protein by phenol:chloroform extraction, the DNA samples were cleaved with piperidine and resolved on 8M urea 12% acrylamide gels. A+G DNA sequencing reactions were carried out with the HF3 probe, and the resulting ladder (lane A+G) was used as the molecular size standard. The diagram at the bottom shows the specific locations of contact points for the bacterially expressed GST-CNR fusion protein in the HF3 site.

coming larger after binding of the antibodies. This phenomena is referred to as a supershift.[39] From the GMSA information can be derived regarding the size of the factor, its tissue distribution, binding affinity and structural requirements. A positive supershift may provide conclusive information as to the identity of the binding protein (figure 5, 6).

Methylation interference: For additional fine tuning of the DNA protein interactions diethyl pyrocarbonate (DEPC) interference (methylation interference) can be used to identify exactly the adenosine and guanine nucleotides involved in protein binding. This chemical variant of DNase I footprinting is based on the DEPC reaction with purines bound to protein generating an alkali-labile abasic site. For this assay DNA is end-labeled, and incubated with either a nuclear extract or purified protein. The DNA-protein complex is treated with DEPC and thereafter the DNA is cut by piperidine at the modified purine nucleotides. The nucleotides bound to protein are recognized by a strong signal after autoradiography of the gel. A sequencing reaction of the same DNA molecule serves as a size marker of this reaction.[40, 41]

In the DNA-protein binding assays presented sofar the DNA and protein are allowed to interact under artificial circumstances *in vitro*. The binding of protein

to DNA can depend on conditions like temperature, pH and salt concentrations. These circumstances are potentially different for each DNA-protein interaction. To identify DNA-protein interactions *in vivo* , the method of methylation interference was modified in order to conserve and recognize the DNA-protein complexes formed in living cells (*in vivo* footprinting). Genomic DNA in cultured cells can be treated by dimethylsulfate, isolated and again the cleavage sites of the methylated DNA can be assessed. For instance a conserved *cis* element (E-box referred to as MLE) on the MLC-2v promoter was studied this way to determine exactly which bases are crucial for the actual DNA-protein interaction (Figure 4). The protein binding to this specific site was later identified as the upstream stimulating factor (USF).[39] The USF is a member of the Helix-Loop-Helix family of transcription factors. The identification of this NABP was possible as previous studies had shown binding of this factor to an identical *cis* element on the tumor suppressor gene p53 promoter.[42]

In vitro *RNA assays:* Tools to study splice events *in vitro* are available as well. As mentioned already the function of spliceosomes can be determined by primer extension. In addition the RNA-protein interaction can be determined in gel mobility shift assays as described for DNA. The affinity of the interaction between protein and RNA can be a determined by cross-linking protein to labeled RNA prior to gel electrophoresis in the presence or absence of competing unlabeled RNA.[36,43] The influence of nucleic acid sequence, as well as the importance of various protein fractions for RNA-protein interaction can be studied.[44]

Mutation analysis: In addition to the *in vitro* and *in vivo* binding studies, functional studies will help to determine the importance of individual *cis* elements. To analyze one specific *cis* element constructs with the normal sequence can be compared to constructs with mutated *cis* element. The inserted mutations vary from one single base, to a complete replacement or deletion of the *cis* element. Oligonucleotides synthesis is used to produce normal and mutated *cis* element containing DNA fragments. These oligonucleotides can be used for GMSA, to study protein-DNA binding, or can be used to generate new constructs harboring mutations by PCR. A regular PCR procedure can be used to mutate the ends of a DNA fragment. By site-directed mutagenesis any position can be altered. For site-directed mutagenesis two PCR fragments are generated with identical mutations and overlapping sequences. The PCR products are mixed, denatured and allowed to reanneal. The overlapping segments of the complementary strands will anneal. The open single strand ends can be filled, and the newly formed DNA template harboring a mutation can be amplified.[45] After cloning the final PCR product in a reporter vector, the effect of the mutation on the level of transcription can be determined through transient transfections using both the normal and mutated sequence. This approach was followed, to study the importance of the HF sites of the rat MLC-2v regulatory sequence[20] and the CRE site on the Kv1.5 regulatory sequence.[44]

Transgenic assays: Data derived from transient transfections form the bases for the construction of transgenic mice to further establish the role of identified DNA-sites in the intact organism.[4,20,46,47] To this end linearized DNA (the promoter-reporter fragment) is injected directly into the pronucleus of a fertilized oocyte. The oocyte can be implanted into the uterus of a pseudo pregnant mouse. If the injected construct is integrated in the genome of the offspring, transgenic lines have been generated. The transgenic experiments outlined here focus on the importance of the modification of *cis* elements of the regulatory sequence of the target gene (table 2). As a reporter gene is part of the transgenic construct, the expression pattern can either be determined histochemically (β-galactosidase: X-gal staining)[24] or by measuring reporter activity biochemically (luciferase: light emission).[20] The transgenics will provide information on tissue specificity of the expression pattern as well as the developmental time frame of transgene expression. Furthermore the effect of pathologic conditions on transgene expression levels can be studied as well. Most often the mouse is used for transgenesis as this model is fast in reproduction, relatively cheap and easy to handle. Surgical techniques can be applied to mimic cardiac disease like myocardial hypertrophy.[47] The limitations of the transgenic technique are due to variable transgene copy number and genomic insertion site. Therefore it is essential to study more then one transgenic mouse line harboring the same construct, to check for a reproducible phenotype.

Cloning of nucleic acid binding factors

Transcription factors are classified either by their trans-acting, protein-protein binding or DNA binding domain. The importance of structural motifs has been shown in Helix-Loop-Helix proteins, leucine zipper, Helix-Turn-Helix (homeodomain) and zinc fingers. These domains are conserved throughout evolution.[48,49] Several different approaches can be taken to identify novel transcription factors.

Degenerate primer PCR: The first strategy is based on knowledge of the DNA sequence of previously cloned NABP. By applying the polymerase chain reaction (PCR)[1] using degenerate primers with homology to NABP coding sequences, new proteins with a conserved domain, similar to the known factor can be identified. Degenerate primers are constructed by making limited changes in the described DNA sequence. These primers will anneal to DNA with sequence homology in a PCR reaction, when the annealing temperature chosen is lower then the melting temperature. The amplified fragment can be subcloned, sequenced and used as a probe to screen a complementary DNA (cDNA) library.[50] With the degenerate primer method different proteins, with a conserved protein domain have been cloned. For example using the homeobox domain cardiac specific homeobox

genes were identified in Xenopus.[51] Similarly by using a zinc-finger domain from the GATA transcription factor family, the chick GATA 6 was cloned.[52] These transcription factors play an important role in contractile protein expression during cardiogenesis. Degenerate primer PCR can also be used to look for cell type or species specific transcription factors. A related method uses low stringency DNA hybridization to identify related nucleotide sequences. This approach has been used by several groups to search for cardiac dominant factors, with homology to the family of myogenic genes. The myogenic transcription factors were first identified in skeletal muscle cells. These factors are considered dominant as they are capable to activate the skeletal gene program in non-muscle cells.[53,54] The introduction of these myogenic factors in cardiac fibroblasts results in skeletal myocyte formation.[55,56] Cardiac NABP with homology to the myogenic factors have been found, and are probably involved in early growth and looping of the heart.[57] However there is no proof yet for a dominant role of these factors in cardiomyocyte differentiation comparable to the skeletal muscle cell.[58]

Affinity chromatography: The second approach exploits the DNA-binding capacity of these proteins to purify the protein or its DNA-binding domain. To purify NABP a method was developed using DNA affinity chromatography. A single or concatenated consensus DNA sequence can be chemically linked to an activated sepharose support column. Nuclear extract harvested from selected tissue or a particular cell line can be applied to the column and the purified proteins can be washed from the sepharose by increasing the salt concentration.[59] The amino acid sequence of part of the purified protein can be determined and used to predict the corresponding DNA sequences. Taking into account the degeneracy of the genetic code a degenerate oligonucleotide can be synthesized for the screening of a cDNA library or for use as a primer for PCR.[60]

Southwestern screening: An alternative method implementing nucleic acid-protein interaction can be used to screen a cDNA expression library for nucleic acid binding domains. To perform this experiment, an expression library may be constructed in which cDNA is inserted into the coding region of the bacteriophage lambda β-galactosidase gene. The phage library can infect a bacterial culture and clones where the DNA binding domain of interest is correctly translated will bind to a labeled nucleic acid probe. For this purpose a concatenated DNA molecule containing a repeat of the DNA-site involved in the DNA-protein interaction, is synthesized. This method is referred to as Southwestern screening, as it involves the interaction between filter-based protein with DNA in the solution (Figure 8).[61] In general only part of the protein, i.e. the DNA-binding domain will be cloned by Southwestern screening. In order to get the full length cDNA of the gene it may be necessary to screen a cDNA library similar to the situation using the method for degenerate primer PCR. This method has been

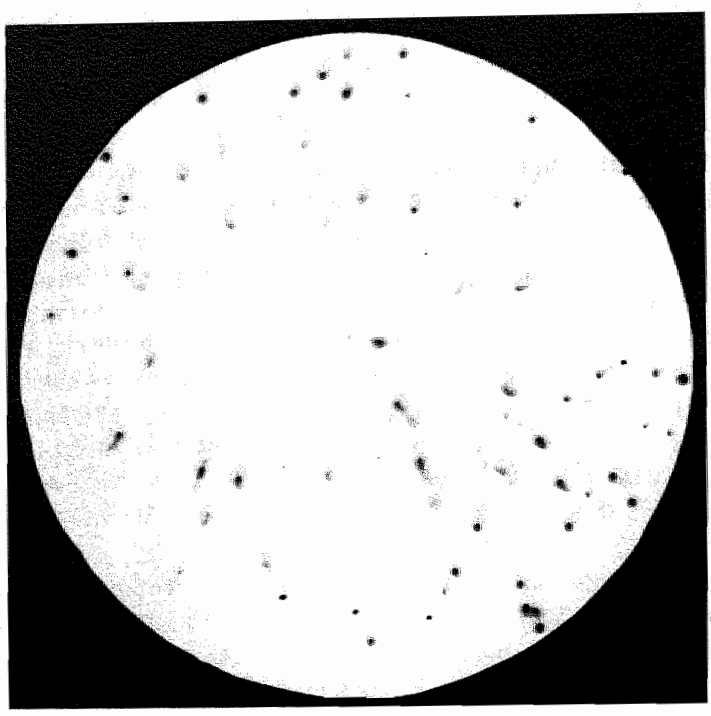


Figure 8. South Western Screening: Second round of South western screening on a neonatal rat λGT 11 expression library. This library expresses cDNA as a galactosidase fusion protein after IPTG induction. For this experiment the conserved HF3 *cis* element of the ventricular myosin light chain 2 was used as a probe. Oligonucleotides containing the HF3 site were end labeled and by ligation a concatenated probe was generated. Oligonucleotide sequence: 5' AGCTTACTGTCTTTT AACCTTGAAGGCA TTTTTA 3'. Picture shows filter containing fusion protein produced by phage infected bacteria. The filter was hybridized with the concatenated probe. Shown is a second round of screening, with several positive clones marked by the black dots. The typical comet appearance of the radioactive signal is caused by the filter lifting procedure.

applied to identify transcription factors interacting with the regulatory sequence the MLC-2v gene like HF3 and HF1B binding proteins.[41,62]

A modification of the Southwestern screening technique involves the use of the purified protein from the DNA-affinity chromatography to generate specific antibodies. Binding of these antibodies to clones of a cDNA expression library can be used to identify bacteriophage clones encoding specific nucleic acid binding domains.[40,63]

Characterization of nucleic acid binding proteins

The cloning of a NABP is one step, the characterization and determination of function will be the next. In general is this the real challenge as function of the NABP is the key issue. The method used to clone the gene may provide some clues as to the anticipated function.

Sequence comparison: Many of the transcription factors identified using the methods described above will be known, or contain domains with high homology to related factors of the same family.[64] This homology can be found in either the DNA-binding, protein-protein interaction or the trans-activation domain. The trans-activation domain of a NABP indicates the domain responsible for the modification of the level of transcription. In addition to sequence comparison, the *in vitro* DNA-protein binding assays like GMSA, will demonstrate the actual DNA binding capacity of the new clone. For these experiments the protein can be produced by *in vitro* transcription translation using a reticulocyte lysate, or by a bacterial expression system. For various reasons it may be worthwhile to generate a bacterial fusion protein. The NABP can be linked for instance to the glutathione S-transferase protein or a poly histidine tag.[65,66] The peptide fused to the NABP allows purification of the bacterial protein. The bacterial fusion protein is useful for GMSA and for the generation of polyclonal antibodies. The antibodies will help to characterize the NABP, as antibody with nuclear extract incubation will give rise to a supershift band in GMSA. Furthermore the antibodies can be used for immunohistochemistry or immunoprecipitation.[67]

Although the knowledge about the regulation of transcription and splicing is accumulating fast, the mechanisms of tissue specific gene expression and cell type specific splicing are still unknown. An intriguing example is the myocyte enhancer factor 2 (MEF2) gene. This NABP plays an important role in skeletal and cardiac muscle specific gene expression. The four genes of MEF2 are expressed in many tissues, however the protein is found predominantly in brain and muscle. The four different genes have numerous splice variants some of which are cell type specific. It is likely that the balance in protein formation of the various splice variants is important for the differentiation pathway of myocytes.[68,69] Clearly the functional characterization of this family of MEF2 proteins is elaborate and time consuming. Up to now most of the knowledge on transcription regulation is obtained from studies involving contractile protein genes. The ongoing integration of cellular electrophysiology with molecular biology will result in ion channel promoter analysis in the years to come. Comparing the sequence of newly identified NABP can help to identify the protein domain responsible for protein-protein interaction, in addition to the domain involved in nucleic acid binding. The information on the structure of the of the NABP can be helpful to assess the nucleic acid binding mechanism. For instance a group of NABP, including HF1B, use zinc ion containing domains for the DNA inter-

action (zinc finger motif).[62] Another group of NABP, the RNA binding proteins often contain a conserved domain: the RNA recognition motif.[70]

Protein structure analysis: Additional structural information can be obtained from crystallized protein analysis and nuclear magnetic resonance spectroscopy. Crystallized protein analysis will further support the characterization of a new NABP and can be helpful in understanding the mechanism of transcription regulation. Advanced computer models allow the reconstruction, using data from structural studies, of three dimensional models indicating the folding and tertiary structure of protein and DNA.[71,72]

Expression pattern: In addition to structural analysis, the expression pattern of the gene can be determined. The expression level in various tissues can be shown by Northern blotting in a semi quantitative manner. The result will permit the identification of tissues with high versus low expression levels. Furthermore by comparing one tissue at different developmental stages, it is possible to judge the onset of expression of a specific NABP for instance in the heart. To acquire information, about which cells contribute to the expression of the gene the mRNA localization within the tissue can be studied by *in situ* hybridization or *in situ* PCR. These methods will show on the cellular level where the gene is expressed (for instance left versus right ventricle) and again by comparing different developmental stages a time course of expression can be determined.[70] These methods to characterize the NABP were applied to study the dynamic expression pattern of the splice factor mSAP49. This protein is expressed at a high level in the heart during a short episode in embryonic development. Its maximal expression coincides with a change in expression of contractile protein isoforms.[8,70,73,74]

Functional assays of nucleic acid binding protein

The structural analysis of the NABP can already provide important clues as to the mechanism of DNA-protein binding. Some proteins bind DNA as monomers, for instance homeodomain containing NABP, whereas others require the binding of a protein partner. For instance the leucine zipper domain is responsible for protein-protein interaction, and transcription factors containing such a domain can only bind to DNA as homo- or heterodimers. Dimerization to activate DNA binding is also reported for NABP belonging to nuclear receptor superfamily. These NABPs bind either as monomer, homodimer or heterodimer.[49,75-77] The fine tuning of transcription often requires post translational modifications of NABP by conformational changes induced either through phosphorylation, or the selective binding of diverse cofactors, like for instance thyroid hormone or retinoic acid.[77-79] These modifications can be overlooked during the structural

Table 2. Strategies to study nucleic acid binding proteins and regulatory DNA *in vivo*

A: Transgene:
- Regulatory DNA analysis: Deletion fragments and wildtype versus mutated cis elements driving a reporter gene.[60]
- NABP overexpression: NABP driven by either non specific or cell type specific promoter.[82]
B: Knock out :
- Underexpression of NABP.[91,92]
C: Knock in:
- Replacement of one allele, where the inserted gene is placed under the control of the regulatory DNA of the exchanged gene.[85]
D: Tissue specific knock out:
- Underexpression model of NABP limited to a specific type of tissue.[87]

analysis. Therefore the role of the various NABP will have to come from functional studies both *in vitro* and *in vivo*.

Functional assays: For the *in vitro* analysis again transient transfection assays can be used. This requires one construct carrying the regulatory element linked to a reporter gene and a second construct containing the NABP cDNA driven by a strong viral promoter, like RSV or CMV promoter. Through co-transfection of both constructs the effect of transient NABP overexpression on the target regulatory DNA can be determined. As a control experiment a known NABP with stimulating effects on the regulatory DNA is used, and as a negative control the new NABP cDNA is omitted. An example of this approach is the study on the effect of MEF-2 overexpression on the activity of a homeobox gene promoter (GAX).[80] To study the role of p21 overexpression on an E2F sites containing promoter-reporter gene construct similar experiments were performed. p21 is believed to be a mediator of p53, a tumor suppressor gene. E2F-1 plays an important role in tumor growth and neointima formation after arterial injury.[81] The outcome of these studies suggest interference of p21 with E2F-1 binding to the E2F DNA element.

Gene targeting: To assess the physiological role of a NABP *in vivo*, the preferred technology would use transgenic or gene targeting strategies to study the effect of overexpression (transgenic), diminished expression (heterozygous knock out: +/-) or no expression (homozygous knock out: -/-) on development and organ (tissue) function (table 2). For these experiments the focus is on the coding region of the NABP, in contrast to the previous application of transgenic technology to study the regulatory DNA sequences. For overexpression studies, a transgene has to be constructed encompassing a selected promoter, which can be either ubiquitously active (viral promoter: RSV, CMV) or tissue specific (flk-1, αMHC, MLC-2v promoter[82-84]). To reduce the level of the NABP, the gene can be destroyed by homologous recombination leading to a reduction of the expression level.

By back crossing a heterozygous knock out line (+/-), homozygous or double knock-out (-/-) off spring can be generated. A variation on these techniques has been developed where the transcribed part of one gene is replaced by another functional gene. In this experiment one copy of gene A is removed and replaced by homologous recombination by gene B. Through this cloning strategy the additional copy of B is placed under the regulatory control of gene A. This knock *in* strategy was successfully applied by London et al.[85], to replace one of the two alleles of the potassium channel Kv1.5 with Kv1.1.

Recent developments in gene targeting technology make ventricular specific knock out of NABP feasible, albeit complex. In order to achieve tissue specific underexpression the target gene named X, has to be isolated and subcloned. In the next step part of the gene for instance exon 2, is placed between two loxP sites. These loxP sites are restriction sites recognized by the enzyme Cre recombinase, and whenever Cre is present it will cut 5' and 3' of exon 2 and ligate the ends of the gene together thereby removing exon 2 on one allele. This will result in diminished levels of X protein. In the absence of Cre however the transcript of X is normal and the protein level will be normal. After manipulation of exon 2 the gene X is re-introduced into the mouse genome through homologous recombination in embryonic stem cells. Simultaneously a transgenic mouse line is generated expressing the Cre recombinase gene in a tissue specific manner, for instance limited to ventricular myocytes.[86] Crossing of the mouse lines will result in reduced levels of protein X, after the tissue specific promoter (α-MHC) driving Cre expression has been activated.[87] The knowledge of NABP and their nucleic acid interactions has triggered the development of new therapeutic strategies, like for example the transfection of myogenic factors (like MyoD) into cardiac fibroblasts. Although no improvement of cardiac function has been shown yet, the induction of myocyte formation was successful.[56] Another strategy used the *cis* element as a decoy to bind functional transcription factors. A repetitive DNA element was made encompassing several binding sites for the E2F protein. The constructed DNA competed with endogenous *cis*-elements for NABP binding and was shown to prevent neointima formation after arterial injury in rat.[88]

The next series of experiments could be aimed at the study of the genomic clone of the NABP to identify transcription start site and splice sites, to identify and characterize the regulatory sequence of the gene. This completes the circle as again we are confronted with a genomic 5' flanking region that will interact with various NABP.

Conclusions

In this paper an outline is given of the most current strategies to study transcription regulation and the proteins that are involved in this complex process. Each

gene, including NABP coding genes themselves, are dependent on transcription activation by binding of NABP to its regulatory DNA. Modern molecular technology permits detailed analysis of DNA-protein interaction. The role of individual bases can be determined together with the essential characteristics of protein domains. Several DNA elements important for cardiac specific gene expression have been identified. Using these *cis* elements, a number of NABP expressed exclusively in the heart, have been isolated and characterized. Through the development of murine gene targeting techniques and by miniaturizing technology, functional studies of NABP genes are feasible. The quest for early tissue specific NABP and dominant cardiac transcription factors will continue. The elucidation of RNA-protein interaction and the mechanism of splicing events is a matter of intensive ongoing research. The ultimate goal of these studies is to understand the mechanism of control of protein production in mammalian cells for research and possibly for therapeutic applications in the near future.

Literature

1. Swynghedauw B, Barrieux A: An introduction to the jargon of molecular biology: Part I. *Cardiovasc Res* 1993;27:1414-1420
2. Swynghedauw B, Moalic JM, Bourgeois F, Barrieux A: An introduction to the jargon of molecular biology. Part II. *Cardiovasc Res* 1993;27:1566-1575
3. van Bilsen M, Chien KR: Growth and hypertrophy of the heart: towards an understanding of cardiac specific and inducible gene expression. *Cardiovasc Res* 1993;27:1140-1149
4. Robbins J: Regulation of cardiac gene expression during development. *Cardiovasc Res* 1996;31:E2-16
5. Vincent CK, Gualberto A, Patel CV, Walsh K: Different regulatory sequences control creatine kinase-M gene expression in directly injected skeletal and cardiac muscle. *Mol Cell Biol* 1993;13:1264-1272
6. Kelly R, Alono S, Tajbakhsh S, Cossu G, Buckingham M: Myosin light chain 3F regulatory sequences confer regionalized cardiac and skeletal muscle specific expression in transgenic mice. *J Cell Biol* 1995;129:383-396
7. Mori Y, Matsubara H, Folco E, Siegel A, Koren G: The transcription of a mammalian voltage-gated potassium channel is regulated by cAMP in a cell-specific manner. *J Biol Chem* 1993;268:26482-26493
8. Lyons GE, Muhlebach S, Moser A, Masood R, Paterson BM, Buckingham M, Perriard JC: Developmental regulation of creatine kinase gene expression by myogenic factors in mouse and chick embryos. *Development* 1991;113:1017-1029
9. Cooper TA, Ordahl CP: A single cardiac Troponin T gene generates embryonic and adult isoforms via developmentally regulated alternative splicing. *J Biol Chem* 1985;11140-11148

10. Jin JP, Lin JJC: Isolation and characterization of cDNA clones encoding embryonic and adult isoforms of rat cardiac Troponin T. *J Biol Chem* 1989;264:14471-14477
11. Nakamura K, Iijima T: Postnatal changes in mRNA expression of the K+ channel in rat cardiac ventricles. *Jpn J Pharmacol* 1994;66:489-492
12. Smale ST, Schmidt MC, Berk AJ, Baltimore D: Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID. *Proc Natl Acad Sci U S A* 1990;87:4509-4513
13. Mullen MP, Smith WJ, Patton JG, Nadal-Ginard B: a-Tropomyosin mutually exclusive exon selection: competition between branchpoint/polypyrimidine tracts determines the default exon choice. *Genes Dev* 1991;5:642-655
14. DeAngelo DJ, DeFalco J, Rybacki L, Childs G: The embryonic enhancer-binding protein SSAP contains a novel DNA-binding domain which has homology to several RNA-binding proteins. *Mol Cell Biol* 1995;15:1254-1264
15. Kamada S, Miwa T: A protein binding to CArG box motifs and to single-stranded DNA functions as a transcriptional repressor. *Gene* 1992;119:229-236
16. Severson WE, Mascolo PL, White MW: Lymphocyte p56L32 is a RNA/DNA-binding protein which interacts with conserved elements of the murine L32 ribosomal protein mRNA. *Eur J Biochem* 1995;229:426-432
17. Rivera-Pomar R, Niessing D, Schmidt-Ott R, Gehring WJ, Jackle H: RNA binding and translational suppression by bicoid. *Nature* 1996;379:746-749
18. Doevendans PA, Ruiz-Lozano P, Bronsaer R, Chien KR, Bilsen van M: Initiation of transcription by the atrial myosin light chain 2 promoter requires the presence of the INR *cis* element. *J Am Coll Cardiol* 1997;29:87-A
19. Zhu H, Garcia AV, Ross RS, Evans SM, Chien KR: A conserved 28-base-pair element (HF-1) in the rat cardiac myosin light-chain-2 gene confers cardiac-specific and alpha-adrenergic-inducible expression in cultured neonatal rat myocardial cells. *Mol Cell Biol* 1991;11:2273-2281
20. Lee KJ, Hickey R, Zhu H, Chien KR: Positive regulatory elements (HF-1a and HF-1b) and a novel negative regulatory element (HF-3) mediate ventricular muscle-specific expression of myosin light-chain 2-luciferase fusion genes in transgenic mice. *Mol Cell Biol* 1994;14:1220-1229
21. Gonzalez GA, Montminy MR: Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. *Cell* 1989;59:675-680
22. Ruiz Lozano P, de Lecea L, Buesa C, Perez de la Osa P, LePage D, Gualberto A, Walsh K, Pons G: The gene encoding rat phosphoglycerate mutase subunit M: cloning and promoter analysis in skeletal muscle cells. *Gene* 1994;147:243-248
23. Yu W, Dahl G, Werner R: The connexin43 gene is responsive to oestrogen. *Proc R Soc Lond B Biol Sci* 1994;255:125-132
24. Ross RS, Navankasattusas S, Harvey RP, Chien KR: An HF-1a/HF-1b/MEF-2 combinatorial element confers cardiac ventricular specificity and established an anterior-posterior gradient of expression. *Development* 1996;122:1799-1809

25. Bergeron D, Barbeau B, Leger C, Rassart E: Experimental bias in the evaluation of the cellular transient expression in DNA co-transfection experiments. *Cell Mol Biol Res* 1995;41:155-159
26. Heim.R, Cubitt.A, Tsien.R: Improved green fluorescence. *Nature* 1995;373:663-664
27. Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC: Green fluorescent protein as a marker for gene expression. *Science* 1994;263:802-805
28. Doevendans PA, Becker DA, An R, Kass RS: The utility of fluorescent *in vivo* reporter genes in molecular cardiology. *Biochem Biophys Res Comm* 1996;222:352-358
29. Heim R, Prasher DC, Tsien RY: Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc Natl Acad Sci U S A* 1994;91:12501-12504
30. Henderson SA, Spencer M, Sen A, Kumar C, Siddiqui MA, Chien KR: Structure, organization, and expression of the rat cardiac myosin light chain-2 gene. *J Biol Chem* 1989;264:18142-18148
31. Simpson P, Shoshana S: Differentiation of rat myocytes in single cell cultures with and without proliferating nonmyocardial cells. *Circ Res* 1982;50:101-116
32. Fuller SJ, Chien KR: Genetic engineering of cardiac muscle cells: in vitro and in vivo. *Genet Eng N Y* 1994;16:17-31
33. Kirshenbaum LA, Schneider MD: Adenovirus E1A represses cardiac gene transcription and reactivates DNA synthesis in ventricular myocytes, via alternative pocket protein- and p300-binding domains. *J Biol Chem* 1995;270:7791-7794
34. Burgess DL, Jones JM, Meisler MH, Noebels JL: Mutation of the Ca2+ channel beta subunit gene Cchb4 is associated with ataxia and seizures in the lethargic (lh) mouse. *Cell* 1997;88:385-392
35. Curran ME, Splawski I, Timothy KW, Vincent GM, Green ED, Keating MT: A molecular basis for cardiac arrhythmia: HERG mutations cause long QT syndrome. *Cell* 1995;80:795-803
36. Valcarcel J, Singh R, Zamore PD, Green MR: The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. *Nature* 1993;362:171-175
37. Schmitz A, Galas DJ: DNase I footprinting: a simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res* 1978;5:3157-3170
38. Wang G, Semenza G: Identification and characterization of transcription factors from mammalian cells, in Adolph KW (ed): *Human molecular genetics*. Orlando, Academic press, 1996, pp 298-320
39. Navankasattusas S, Sawadogo M, van Bilsen M, Dang CV, Chien KR: The basic Helix-Loop-Helix protein upstream stimulating factor regulates the cardiac ventricular myosin light-chain 2 gene via independent cis regulatory elements. *Mol Cell Biol* 1994;14:7331-7339
40. Sturm RA, Das G, Herr W: The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. *Genes Dev* 1988;2:1582-1599
41. Ruiz-Lozano P, Doevendans PA, Brown A, Chien KR: A *sex-lethal* related factor (CNR) is a negative regulator of MLC-2v expression. *Circulation* 1995;92:755-A

42. Reisman D, Rotter V: The Helix-Loop-Helix containing transcription factor USF binds to and transactivates the promoter of the p53 tumor suppressor gene. *Nucleic Acids Res* 1993;21:345-350
43. Siomi H, Siomi MC, Nussbaum RL, Dreyfuss G: The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein. *Cell* 1993;74:291-298
44. Patton JG, Porro EB, Galceran J, Tempst P, Nadal Ginard B: Cloning and characterization of PSF, a novel pre-mRNA splicing factor. *Genes Dev* 1993;7:393-406
45. Higuchi R: Using PCR to engineer DNA, in Erlich HA (ed): *PCR technology.* New York, W.H. Freeman and company, 1992, pp 61-70
46. Doevendans PA, Hunter JJ, Lembo G, Wollert KC, Chien KR: Strategies for studying cardiovascular diseases in transgenic mice and gene-targeted mice, in Monastersky GM, Robl JM (eds): *Strategies in transgenic animal science.* Washington, American Society For Microbiology, 1995, pp 107-144
47. Kubalak S, Doevendans PA, Rockman H, Hunter JJ, Tanaka N, Ross J, Chien KR: Molecular analysis of cardiac muscle diseases based on mouse genetics, in Adolph KW (ed): *Human molecular genetics.* Orlando, Academic Press, 1996, pp 470-487
48. Doevendans PA, van Bilsen M: Transcription factors and the cardiac gene programme. *Int J Biochem & Cell Biol* 1996;28:387-403
49. Papavassiliou A: Molecular medicine: transcription factors. *N Eng J Med* 1995;322:45-47
50. Compton T: Degenerate primers for DNA amplification, in Innes MA (ed): *PCR protocols: A guide to methods and applications.* 1990, pp 39-45
51. Evans SM, Yan W, Murillo MP, Ponce J, Papalopulu N: tinman, a Drosophila homeobox gene required for heart and visceral mesoderm specification, may be represented by a family of genes in vertebrates: XNkx-2.3, a second vertebrate homologue of tinman. *Development* 1995;121:3889-3899
52. Laverriere AC, MacNeill C, Mueller C, Poelmann RE, Burch JB, Evans T: GATA-4/5/6, a subfamily of three transcription factors transcribed in developing heart and gut. *J* Biol Chem 1994;269:23177-23184
53. Weintraub H, Davis R, Tapscott S, Thayer M, Krause M, Benezra R, Blackwell T, Rupp R, Hollenberg S, Huang Y, Lassar AB: The MyoD family: nodal point during specification of the muscle cell lineage. *Science* 1989;251:761-766
54. Lossar A, Paterson BM, Weintraub H: Transfection of a cDNA locus that mediates the conversion of 10T1/2 fibroblasts to myoblasts. *Cell* 1986;47:649-655
55. Murry CE, Kay MA, Bartosek T, Hauschka SD, Schwartz SM: Muscle differentiation during repair of myocardial necrosis in rats via gene transfer with MyoD. *J Clin Invest* 1996;98:2209-2217
56. Tam SK, Gu W, Nadal Ginard B: Molecular cardiomyoplasty: potential cardiac gene therapy for chronic heart failure. *J Thorac Cardiovasc Surg* 1995;109:918-923
57. Srivastava D, Cserjesi P, Olson EN: A subclass of bHLH proteins required for cardiac morphogenesis. *Science* 1995;270:1995-1999

58. Olson EN, Srivastava D: Molecular pathways controlling heart development. *Science* 1996;272:671-676
59. Kadonga JT, Tjian R: Affinity purification of sequence-specific DNA binding proteins. *Proc Natl Acad Sci U S A* 1986;83:5889-5893
60. Norman C, Runswick M, Pollock R: Isolation and properties of cDNA clones encoding SRF, a transcription factor that binds to the c-fos serum response element. *Cell* 1988;55:989-1003
61. Vinson CR, LeMarco K, Johnson P, Landschulz WH, McKnight SL: In situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. *Genes Dev* 1988;2:801-806
62. Zhu H, Nguyen VT, Brown AB, Pourhosseini A, Garcia AV, van Bilsen M, Chien KR: A novel, tissue-restricted zinc finger protein (HF-1b) binds to the cardiac regulatory element (HF-1b/MEF-2) in the rat myosin light-chain 2 gene. *Mol Cell Biol* 1993;13:4432-4444
63. DeAngelo DJ, DeFalco J, Childs G: Purification and characterization of the stage-specific embryonic enhancer-binding protein SSAP-1. *Mol Cell Biol* 1993;13:1746-1758
64. Yu YT, Breitbart RE, Smoot LB, Lee Y, Mahdavi V, Nadal Ginard B: Human myocyte-specific enhancer factor 2 comprises a group of tissue-restricted MADS box transcription factors. *Genes Dev* 1992;6:1783-1798
65. Van Dyke MW, Sirito M, Sawadogo M: Single-step purification of bacterially expressed polypeptides containing an oligo-histidine domain. *Gene* 1992;111:99-104
66. Hakes DJ, Dixon JE: New vectors for high level expression of recombinant proteins in bacteria. *Anal Biochem* 1992;202:293-298
67. Sawadogo M, Van Dyke MW: Indirect use of immobilized metal affinity chromatography for isolation and characterization of protein partners. *Genet Eng N Y* 1995;17:53-65
68. Kaushal S, Schneider JW, Nadal Ginard B, Mahdavi V: Activation of the myogenic lineage by MEF2A, a factor that induces and cooperates with MyoD. *Science* 1994;266:1236-1240
69. Martin JF, Miano JM, Hustad CM, Copeland NG, Jenkins NA, Olson EN: A Mef2 gene that generates a muscle-specific isoform via alternative mRNA splicing. *Mol Cell Biol* 1994;14:1647-1656
70. Ruiz-Lozano P, Doevendans PA, Brown A, Gruber PJ, Chien KR: Developmental and tissue restricted expression of a murine spliceosome related protein mSAP49 during cardiogenesis. *Dev Dyn* 1997;208:482-490
71. Geiger JH, Hahn S, Lee S, Sigler PB: Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science* 1996;272:830-836
72. Arnold R, Burcin M, Kaiser B, Muller M, Renkawitz R: DNA bending by the silencer protein NeP1 is modulated by TR and RXR. *Nucleic Acids Res* 1996;24:2640-2647
73. Gorza L, Ausoni S, Merciai N, Hastings KE, Schiaffino S: Regional differences in troponin I isoform switching during rat heart development. *Dev Biol* 1993;156:253-264

74. Lyons GE, Schiaffino S, Sassoon D, Barton P, Buckingham ME: Developmental regulation of myosin gene expression in mouse cardiac cells. *J Cell Biol* 1990;111:2427-2436
75. Neyses L, Vetter H: Molecular biology of oncogenes and cardiovascular hypertrophy. *J Hypertens* 1992;10:1447-1452
76. Kliewer SA, Umesono K, Mangelsdorf DJ, Evans RM: Retinoid X receptor interacts with nuclear receptors in retinoic acid, thyroid hormone and vitamin D3 signalling. *Nature* 1992;355:446-449
77. Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schutz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P, et al: The nuclear receptor superfamily: the second decade. *Cell* 1995;83:835-839
78. Morkin E: Regulation of myosin heavy chain genes in the heart. *Circulation* 1993;87:1451-1460
79. Sorger PK, Lewis MJ, Pelham HR: Heat shock factor is regulated differently in yeast and Hela cells. *Nature* 1987;329:81-84
80. Andres V, Fisher S, Wearsch P, Walsh K: Regulation of Gax homeobox gene transcription by a combination of positive factors including myocyte-specific enhancer factor 2. *Mol Cell Biol* 1995;15:4272-4281
81. Shiyanov P, Bagchi S, Adami G, Kokontis J, Hay N, Arroyo M, Morozov A, Raychaudhuri P: p21 Disrupts the interaction between cdk2 and the E2F-p130 complex. *Mol Cell Biol* 1996;16:737-744
82. Hunter JJ, Tanaka N, Rockman HA, Ross J, Jr., Chien KR: Ventricular expression of a MLC-2v-ras fusion gene induces cardiac hypertrophy and selective diastolic dysfunction in transgenic mice. *J Biol Chem* 1995;270:23173-23178
83. Ronicke V, Risau W, Breier G: Characterization of the murine FLK-1 promoter. *Circ Res* 1996;79:277-285
84. LI L, Miano JM, Mercer B, Olson EN: Expression of the SM22alpha promoter in transgenic mice provides evidence for distinct transcriptional regulatory programs in vascular and visceral smooth muscle cells. *J Cell Biol* 1996;132:849-859
85. London B, Hill J, Nguyen H, Schieferl S, Nadal Ginard B: In-vivo mutagenesis of the mouse heart: targeted replacement of the murine delayed rectifier potassium channel mKv1.5 with the rat brain delayed rectifier channel rKv1.1. *Circulation* 1995;92:I-195-A
86. Gu H, Marth JD, Orban PC, Mossmann H, Rajewsky K: Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. *Science* 1994;265:103-106
87. Frenkel PA, Agah R, French BA, Michael LH, Overbeek PA, Schneider MD: Targeted expression of Cre recombinase provokes cardiac-restricted site spedific gene rearrangement in ventricular muscle in vivo. *J Am Coll Cardiol* 1997;29:3A

88. Morishita R, Gibbons GH, Horiuchi M, Ellison KE, Nakama M, Zhang L, Kaneda Y, Ogihara T, Dzau VJ: A gene therapy strategy using a transcription factor decoy of the E2F binding site inhibits smooth muscle proliferation in vivo. *Proc Natl Acad Sci U S A* 1995;92:5855-5859
89. Pollock R, Treisman R: Human SRF-related proteins: DNA-binding properties and potential regulatory targets. *Genes Dev* 1991;5:2327-2341
90. Hasty P, Bradley A, Morris JH, Edmondson DG, Venuti JM, Olson EN, Klein WH: Muscle deficiency and neonatal death in mice with a targeted mutation in the myogenin gene. *Nature* 1993;364:501-506
91. Dyson E, Sucov HM, Kubalak SW, Schmid Schonbein GW, DeLano FA, Evans RM, Ross J, Jr., Chien KR: Atrial-like phenotype is associated with embryonic ventricular failure in retinoid X receptor alpha -/- mice. *Proc Natl Acad Sci U S A* 1995;92:7386-7390
92. Lyons I, Parsons LM, Hartley L, Li R, Andrews JE, Robb L, Harvey RP: Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene Nkx2-5. *Genes Dev* 1995;9:1654-1666
93. Sanger F, Coulson AR: A rapid method for determining sequences in DNA by primed synthesis and DNA polymerase. *J Mol Biol* 1975;94:444-448
94. Strauss EC, Andrews NC, Higgs DR, Orkin SH: In vivo footprinting of the human alpha-globin locus upstream regulatory element by guanine and adenine ligation-mediated polymerase chain reaction. *Mol Cell Biol* 1992;12:2135-2142

Chapter

Developmental expression of the murine spliceosome associated protein mSAP49

Pilar Ruiz-Lozano, Pieter Doevendans, Anne Brown, Peter J. Gruber and Kenneth R. Chien

Published: in Development Dynamics 1997; 208: 482-490

Abstract

Background: In the process of identifying nucleic acid binding proteins interacting with a conserved *cis* element (HF3) of the ventricular myosin light chain 2 promoter, the mouse homologue of the human spliceosome associated protein SAP49, mSAP49 was cloned. Here we report the sequence of cDNA and protein of mSAP49, together with the expression pattern of this gene during mouse embryonic development.

Methods: Southwestern screening was performed on a rat ventricular cDNA expression library using the HF3 element as a probe. The identified clones were analyzed by sequencing and sequence comparison and used to obtain the complete mSAP49 cDNA. The expression pattern of mSAP49 was studied during development by Northern blotting, rt-PCR and whole mount in *situ* hybridization.

Results: mSAP49 contains two RNA recognition motifs in the N-terminus of the predicted amino-acid sequence, and a highly basic C-terminus rich in glycine/proline. mSAP49 displayed stage dependent expression during cardiac development. In the adult mouse, mSAP49 is widely distributed, although it was found at relatively low levels in the heart. *In situ* hybridization analysis of mSAP49 mRNA distribution in staged mouse embryos showed that mSAP onset occurs later in the heart than in other embryonic tissues. While mSAP49 expression was found at day 10.0 pc in the optic eminence, optic vesicle, hindbrain, and somites, it was not present in cardiac tissue. mSAP49 was detected in the ventricles at day 11.5, and at day 13.5 it was also detected in the atria. Northern analysis showed that mSAP49 mRNA displayed a peak of expression in the heart at days 14.0-15.0 pc and its abundance decayed thereafter.

Conclusions: This dynamic pattern of cardiac expression suggests that mSAP49 is contributing to a change in the ratio of spliceosome components during cardiac growth and development, which could have consequences for tissue-specific splicing or RNA stabilization.

Introduction

During vertebrate heart development, the expression of contractile proteins is regulated at both the transcriptional and post-transcriptional levels. For example, during avian and mammalian cardiogenesis, the expression of a single troponin T (cTnT) gene increases and concomitantly the cTnT protein switches from the embryonic to adult isoform.[1,2] The mechanism for this switch is known to involve a developmentally regulated alternatively spliced transcript.[3] Similarly, the expression of a vast array of tropomyosin isoforms during cardiac development, and the switch from non-muscle to muscle isoforms, is also controlled at the level of pre-mRNA splicing.[4] In addition, the developmentally regulated switch in contractile protein isoforms appears to correlate with developmental changes in cardiac function.[5]

At present, relatively little is known regarding developmental changes in spliceosome associated proteins that might mediate these switches during the course of cardiogenesis. The process of pre-mRNA splicing requires the concerted actions of both *cis*- and *trans*-acting factors localized in the spliceosome, a ribonucleoprotein (RNP) complex comprised of small nuclear RNPs (snRNPs) U1, U2, U4, U5 and U6 and a number of non-snRNP factors (for a review, Moore et al.[6]) of which only a few have been characterized to date. In alternatively spliced transcripts, splice recognition sites that are efficiently used in one cell type may be overlooked by the splicing machinery in other cell types. Therefore, regulated alternative pre-mRNA processing must involve either differences in *cis*- and *trans*-acting factors or the existence of additional tissue-specific factors. Splice site choice may not always require the presence of specific factors but can be mediated by differences in the activities or ratios between general splicing factors.[7-9]

In the course of screening for proteins that bind to the HF-3 negative regulatory element in the ventricular myosin light chain-2 (MLC-2v)[10,11] we have cloned the murine homologue of the human spliceosome-associated protein SAP49, mSAP49. Spliceosome-associated protein SAP49 is specifically associated with U2 snRNP, interacts with pre-mRNA in a region immediately upstream of the branchpoint sequence in the prespliceosomal complex, and also directly interacts with SAP145, another U2 snRNP associated protein. In this study, we show that mSAP49 transcript accumulation in the heart is temporally regulated during the course of cardiogenesis, thereby suggesting its potential involvement in the developmental regulation of spliceosome activity in the heart.

Methods

Screening of cDNA libraries. To isolate factors that bind to the HF3 site in the MLC-2v promoter, $2x10^6$ pfu λgt11 from a rat neonatal ventricle cDNA expression library[12] were screened with a 5' end labeled HF-3 oligo 5'AGCTTCCACTGTCTCTTTAACCTTGAAGCATTTTTA 3', from positions -211 to -176 of the rat MLC-2v promoter.[10] To obtain a double stranded probe, the sense strand was end labelled and hybridized with a 10-fold excess of cold antisense strand by heating to 94°C for 5 min. and gradually cooling to room temperature. Subsequently, the Hind III cohesive ends of the probe were ligated using T4 DNA ligase (GibcoBRL, Gaithersburg MD). The probe was purified over a G25 Sepharose column and its size evaluated by agarose gel electrophoresis. Southwestern screening of the rat neonatal heart λgt11 cDNA library was performed as described elsewhere.[13] For sequence analysis, positive clones were plaque purified and subcloned into the Not I site of pBluescript II SK(+) (Stratagene, La Jolla CA). To isolate full-length cDNAs, conventional DNA/DNA hybridization was performed using the primary cDNA clone to probe a mouse 14-day embryoid body cDNA library in λgt11 (kindly provided by Dr. J. Robbins).

Northern blotting analyses. Two and a half micrograms of polyA$^+$ RNA from different adult mouse tissues: liver, spleen, skeletal muscle, kidney, brain and heart (Clonetech, Palo Alto CA) or 30μg of total RNA from a pool embryos, were separated on a formamide agarose gel, transferred to a nylon filter (N-Hybond, Amersham, England), fixed by UV cross-linking, and hybridized with $3x10^6$cpm/ml of the ^{32}P random labeled 1.5Kb full length Not I fragment, following manufacturer's directions (Amersham). Washes were performed under moderate stringency (0.2x SSC, 65°C). Densitometry analysis of the blots was performed using the IS-1000Digital Imaging System (Alpha Innotech Corporation, San Leandro CA).

rt-PCR. In vitro differentiation of ES cells, polymerase chain reaction and analysis of the amplified products were performed as previously described. [14]

In situ hybridization. The full length mSAP49 was cloned into pBluescript SKII+ (Stratagene), cut with Hae II. A FITC-labeled cRNA probe was generated by *in vitro* transcription with T7 RNA polymerase and the 10x DIG RNA labeling mixture obtained from Boehringer Mannheim (Indianapolis, Il), to yield an antisense riboprobe of 373 nucleotides (nt). Whole mount *in situ* hybridization was performed using a modification of the procedure described by Christiansen et al.[15] Specimens were photographed submerged in PBS on a 1% agarose petri dish with a stereomicroscope (Zeiss SV6) and slides were digitized. Figures were composed in Adobe Photoshop 3.0, QuarkXPress 3.31, and printed by Fuji Pictography.

Results and discussion

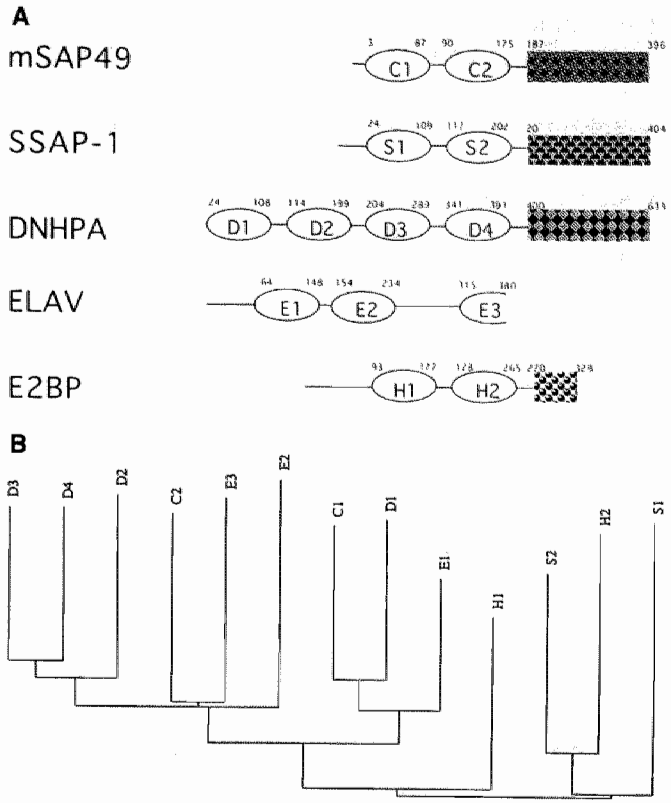
Isolation of mSAP49 cDNAs and predicted protein sequence

In the search for factors which might interact with the HF-3 element in the ventricular myosin light chain promoter[10], we isolated a murine homologue of the human spliceosome associated factor SAP49.[16] A stop codon (position 1280 nt), a putative polyadenylation signal (1485 nt), and a polyA tract were identified, identifying the 3' end of the mRNA (figure 1). The complete mSAP49 cDNA was 1,515 bp long, with a 1,272 bp open reading frame and a predicted 424 amino-acid protein product (figure 1). The predicted polypeptide contained a C-terminal region that was both glycine (16%) and proline rich (34%) and highly basic, with a predicted isoelectric point of 12.18. The N-terminus contained two RNA recognition motifs (RRMs[17]: underlined in figure 1, and figure 2A). Predicted polypeptides of mouse and human SAP49 displayed a 100% conservation within the RRM domains and shared a high glycine/proline content in the C-terminal region; however the sequence in this domain was not strictly conserved, with a 47% divergence in a 71 amino-acid stretch (bold in figure 1). Whether this divergence represents an alternatively spliced form or interspecies variability is unknown.

RRM containing proteins have one or more copies of the RRM domain in the N-terminus and a highly basic auxiliary domain in the C-terminus. These proteins were initially reported as RNA-binding proteins, but recently alternate roles, such as DNA binding[18] and/or protein/protein interactions[16], have been postulated for this protein family. We compared the mSAP49 predicted polypeptide sequence with similar proteins in the data bank, in search for a structure/function relationship. Two proteins involved in transcriptional regulation, the Stage Specific Activator 1 from sea urchin and the human liver-specific protein E2BP[18,19], share two RRM domains and a highly basic C-terminus with mSAP49. Other RRM containing proteins with structural similarities to mSAP49 include the human polyA-Binding protein[20] which has four RRM repeats and the mammalian Elav homolog[21] that has two complete and one incomplete RRM repeat with no accessory C-terminal sequences (figure 2A). Amino-acid sequences from the entire RRM domain of mSAP49 were aligned to closely related RRM domains (figure 2C) using a PILEUP program (GCG, University of Wisconsin) and a dendrogram was constructed representing the relat ionship among RRM domains, prepared from a multiple alignment.[22] From these considerations, it became apparent that the N-terminal RRM domain in the mSAP49 polypeptide (C1, figure 2A) is closest to the N-terminal RRM in the human polyA binding protein (D1, figure 2A) and the C2 domain is closest to D2 in a manner that seems to be dependent upon their position in the polypeptide sequence. If the criterion followed is the number of repeats, then multifunctional proteins with RNA and

```
TTTCGCCATGGCTGCCGGACCGATCTCCGAACGGAATCAGGATGCCACGGTGTACGTG   58
        M  A  A  G  P  I  S  E  R  N  Q  D  A  T  V  Y  V      17

GGAGGTCTAGACGAGAAAGTGAGCGAGCCACTGCTATGGGAGCTCTTTCTCCAGGCAGGG 118
G  G  L  D  E  K  V  S  E  P  L  L  W  E  L  F  L  Q  A  G    37

CCAGTGGTCAACACCCACATGCCCAAGGACAGAGTCACTGGCCAGCACCAGGGCTATGGC 178
P  V  V  N  T  H  M  P  K  D  R  V  T  G  Q  H  Q  G  Y  G    57

TTTGTTGAATTCCTGAGCGAGGAAGATGCCGACTATGCCATTAAGATTATGAACATGATC 238
F  V  E  F  L  S  E  E  D  A  D  Y  A  I  K  I  M  N  M  I    77

AAACTCTATGGAAAGCCAATACGGGTGAACAAGGCCTCAGCTCACAACAAAAACCTGGAT 298
K  L  Y  G  K  P  I  R  V  N  K  A  S  A  H  N  K  N  L  D    97

GTGGGAGCCAACATTTTCATTGGAAATCTGGACCCAGAAATTGATGAAAAGCTGCTTTAC 358
V  G  A  N  I  F  I  G  N  L  D  P  E  I  D  E  K  L  L  Y   117

GATACTTTCAGCGCCTTTGGAGTCATCCTACAGACCCCCAAGATCATGCGGGACCCTGAC 418
D  T  F  S  A  F  G  V  I  L  Q  T  P  K  I  M  R  D  P  D   137

ACAGGCAACTCCAAGGGTTACGCCTTCATTAATTTTGCCTCCTTCGATGCTTCGGATGCA 479
T  G  N  S  K  G  Y  A  F  I  N  F  A  S  F  D  A  S  D  A   157

GCAATTGAGGCCATGAACGGGCAGTACCTGTGTAACCGTCCTATCACTGTGTCTTATGCC 538
A  I  E  A  M  N  G  Q  Y  L  C  N  R  P  I  T  V  S  Y  A   177

TTCAAGAAGGACTCTAAGGGTGAACGCCATGGCTCAGCAGCTGAAAGACTCCTGGCAGCC 598
F  K  K  D  S  K  G  E  R  H  G  S  A  A  E  R  L  L  A  A   197

CAGAACCCGGCTGTCCCAGCTGACCGCCCTCACCAGCTGTTTGCCGATGCACCCCCTCCG 658
Q  N  P  A  V  P  A  D  R  P  H  Q  L  F  A  D  A  P  P  P   217

CCCTCTGCCCCCAATCCCGTGGTTTCATCCCTGGGTTCTGGGCTTCCTCCACCAGGCATG 718
P  S  A  P  N  P  V  V  S  S  L  G  S  G  L  P  P  P  G  M   237

CCGCCTCCTGGCTCTTTTCCACCTCCAGTGCCACCTCCTGGGGCCCTCCCTCCTGGGATT 778
P  P  P  G  S  F  P  P  P  V  P  P  P  G  A  L  P  P  G  I   257

CCCCCAGCCATGCCCCCACCACCTATGCCACCTGGGGCTGGAGGACATGGCCCCCCATGC 838
P  P  A  M  P  P  P  P  M  P  P  G  A  G  G  H  G  P  P  C   277

AGCAGGAACTCCAGGGGCTGGACATCCTGGTCACGGACATTCACATCCTCATCCATTCCA 898
S  R  N  S  R  G  W  T  S  W  S  R  T  F  T  S  S  S  I  P   297

CCAGGTGGGATGCCCCATCCAGGGATGTCCCAGATGCAGCTGGCCCACCATGGCCCCCAT 958
P  G  G  M  P  H  P  G  M  S  Q  M  Q  L  A  H  H  G  P  H   317

GGCCTAGGACACCCCCATGCTGGGCTCCGGGCTCTGGGGGGCAGCCACCACCTCGGCCAC1018
G  L  G  H  P  H  A  G  L  R  A  L  G  G  S  H  H  L  G  H   337

CTCCTGGAATGCCTCCATCCTGGACCTCCTCCAATGGGCATGCCCCCCCGAGGGCCTCCT1078
L  L  E  C  L  H  P  G  P  P  P  M  G  M  P  P  R  G  P  P   357

TTTGGATCTCCCATGGGTCACCCAGGTCCCATGCCTCCACACGGCATGCGTGGGCCTCCT1138
F  G  S  P  M  G  H  P  G  P  M  P  P  H  G  M  R  G  P  P   377

CCATTGATGCCCCCTCATGGATACACGGGTCCTCCAAGACCCCCTCCCTATGGCTACCAG1198
P  L  M  P  P  H  G  Y  T  G  P  P  R  P  P  P  Y  G  Y  Q   397

CGGGGGCCCCTCCCGCCACCCAGACCCACTCCACGGCCCCCAGTTCCACCTCGTGGTCCT1258
R  G  P  L  P  P  P  R  P  T  P  R  P  P  V  P  P  R  G  P   417

CTTCGGGGCCCACTTCCTCAGTAAATTGCCATCCTTTGTCTCCTCCTTTTATCCTCCCTT1318
L  R  G  P  L  P  Q  *                                       424

TATCTTCATTTCTTGGACCAATCAGAGATGCTGTAGCTTGTCAAGGCTAATGTACTGACC1368

CTTCTCAGTAACTCTGTCCCCATTCCAATGTCATGTTTTCCACTGGGGATTTTCTTCATT1438

TTTATTTTAAAATGTTGGTCCTAATTTCTTTACACATGTATAGAAAAATAAACTATGCT1498

TCTTGTTTTAAAAAAAA
                                                             1515
```

Figure 1. mSAP cDNA nucleotide sequence and predicted protein sequence. The two RNA recognition motifs domains are underlined. An in-frame stop codon is marked by*. The putative polyadenylation signal at the 3' end of the transcript is also underlined.



C

```
c1  ELFLQAGPVVNT HMPKDRVTGQHQGYGFVEFLSEEDADYAI KIMNMIKLYG  KPIRVN   KASAHNKNLDVGANIFIGNLDP EIDEKL
c2  DTFSAFGVILQTPKIMRDPDTGNSKGYAFINFASFDASDAAI EAMNGQYLCN  RPITVSYAFKKDSKGERHGSAAERLLAAQNP AVPADR
s2  ELFRSHGNV EDVHIMNDTDTGKHRGFGFVTLDSEEAVEKLV R MHHLELKG  KSMEIK   KAQPKMNRGFGGPGGQGGPGGPGGFPQGG
e1  SLFGSIGEI ESCKLVRDKITGQSLGYGFVNYIDPKDAEKAI NTLNGLRLQT  KTIKVSYARPSSASIR    DANLYVSGL PKTMTQKE
e2  QLFSQYGRII TSRILVDQVTGVSRGVGFIRFDKRIEAEEAI KGLNGQKPSGATEPITVKFANNPSQKSS   QALLSQLYQSPNRRYP
e3  QLFGPFGAV NNVKVIRDFNTNKCKGFGFVTMTNYDEAAMAI ASLNGYRLGD  RVLQVSFKTNKAHKS
d1  EKFSPAGPIL SIRVCRDMITRRSLGYAYVNFQQPADAERAL DTMNFDVIKG  KPVRIMWSQRDPSLRK  SGVGNIFIKNLD KSIDNK
d2  DTFSAFGNIL SCKVVCD  ENGSKGYGFVHFETQEAAERAI EKMNGMLLND  RKVFVGRFKSRKEREAELGARAKEFTNV   YIKNFG
d3  RLKDLFGPAL SVKVMTD ESGKSKGFGFVSFERHEDAQKAV DEMNGKELNG  KQIYVGRAQKKVERQT    ELKRKFEQMKQDRITRYQ
d4  KEFSPFGTI TSAKVMME  GGRSKGFGFVCFSSPEEATKAV TEMNGRIVAT  KPLYVALAQRKEERQAHL   TNQYMQRMASVRAVPN
h1  DYFSKFGEVV DCTLKLDPITGRSRGFGFVLFKESESVDKVM DQKE HKLNG  KVIDPKRAKAMKTKEP    VKKIFVGGLSPDTPEEKI
h2  EYFGGFGEV ESIELPMDNKTNKRRGLCFITFKEEEPVKKIMEKKYHNVGLSK  CEIKVAMSKEQYQQQQQ WGSRGGFAGRARGRGGDQ
s1  AYFEKFGKLSDIILMMDKDKPGONKGFGFVTFADPACVDDVT NEKN HNLEG  KGLDCKRCKARGSEKR  MGPGDQRTKKVFVGGISQQ
```

Figure 2. (A) Schematic representation of the mSAP49 predicted protein and comparison with other related gene products. SSAP-1: Stage Specific Activator 1.[18] DNHPA: Human poly-A binding protein[20] ELAV: Human *elav* [21]. Liver-specific E2BP [19] The RRM domains are represented in circles. The auxiliary C-terminal domains are boxed. Numbers indicate amino-acid position. (B) Dendogram that represents the relationship among RRM domains prepared from a multiple alignment [22] and summarizes the similarity among the RRM containing proteins by sequence. (C) Sequence comparison of RRM domains, from mSAP49 and other related proteins: the two RRM domains in mSAP49 (C1 and C2), in SSAP-1 (S1 and S2), in E2BP (H1 and H2), the four domains in DNHPA (D1-D4) and the three domains in Human Elav (E1-E3) are aligned. Numbers indicate the position of the domain in the protein, where 1 is positioned at the N-terminus of the RRM domains within an amino-acid sequence.

DNA binding activities are composed of two RRM domains, while RNA binding proteins may have additional domains.

The fact that we cloned mSAP49 on the basis of its ability to bind to DNA sequences, and previous results with human SAP49, suggests that mSAP49 may be multifunctional protein. Previous reports have indicated multifunctional roles for several nucleic acid binding factors. For example, the RNA binding protein YB1 binds to a cardiac-specific promoter and positively regulates its transcription.[13] In addition, the homeobox containing protein *bicoid*, initially described as a DNA binding protein, has been shown to bind to *caudal* (cad) mRNA to control cad translation in a region-specific manner.[23]

mSAP49 expression is regulated during mouse development

To determine the tissue distribution of mSAP49 mRNA, we performed Northern blot analyses with poly A^+ RNA derived from various adult mouse tissues: liver, skeletal muscle, spleen, kidney, brain, and heart. The mSAP49 mRNA (a single 1.5 Kb band) was widely distributed, although it was found at relatively low levels in heart, as confirmed after normalization with a GAPDH probe (figure 3A). To further establish that this differential pattern of expression was not simply a function of the quality of the cardiac RNA, a cDNA encoding MLC-2v, a marker for ventricular muscle, was hybridized with the same Northern blot. A specific MLC-2v signal was present in the heart and at lower levels in skeletal muscle (figure 3A), consistent with the expression pattern of MLC-2v in adult tissues previously described.[24]

Embryonic stem cells (ES cells) can aggregate and form spontaneously beating embryoid bodies (EB) which sequentially express cardiac-specific markers and serve as a model system to study cardiogenesis.[14] To determine the onset of appearance of the mSAP49 message, a semi-quantitative reverse transcriptase polymerase chain reaction (rt-PCR) was performed at sequential days during EB development. An mSAP49-specific band was found in both fibroblasts and ES cells, and no significant changes in expression were detected following the differentiation of ES cells into beating EB (figure 3B). These results suggest that mSAP49 onset is very early and that its expression might not be regulated. Alternatively, they may represent a failure of the EB system to detect developmental modulation of non-cardiac-specific genes. In Northern-blot experiments using total RNA from isolated embryonic hearts at days 14-15 days post coitum (pc) and adult heart, mSAP49 was expressed at higher levels in cardiac embryonic tissue relative to adult heart (figure 3C), thus suggesting that mSAP mRNA content varies with cardiac development.

Subsequently, we studied mSAP49 mRNA expression in the developing mouse by whole mount *in situ* hybridization utilizing staged embryos at early days of development. We did not detect mSAP49 mRNA in the mouse embryo at days

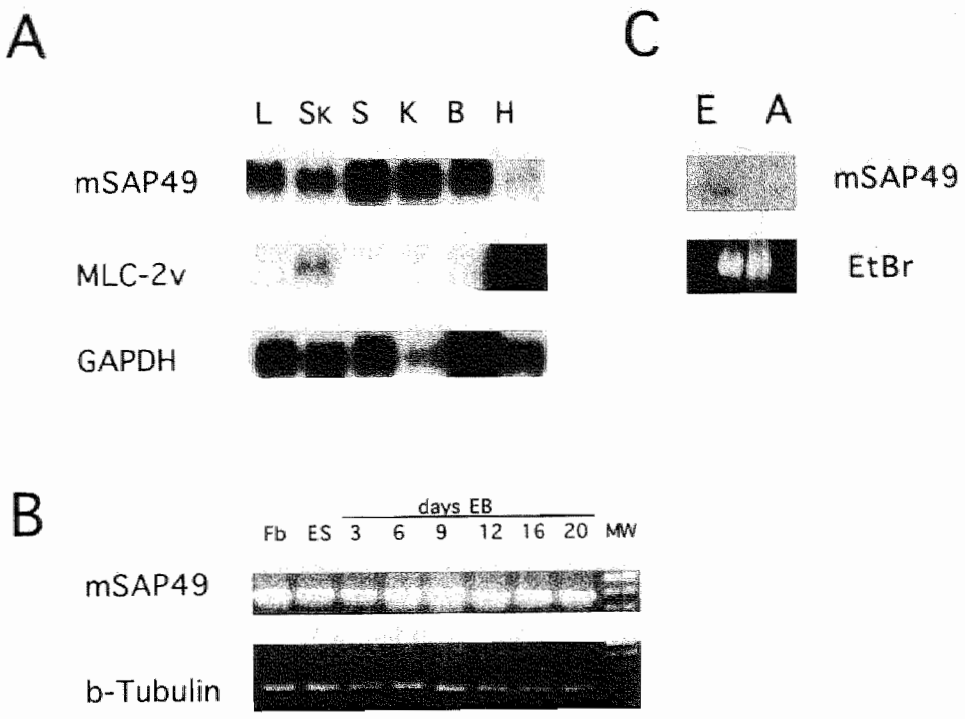


Figure 3. (A) Northern blot using 2.5 µg polyA⁺ mRNA per lane from adult mouse tissues. Right to left: Liver, Skeletal muscle, Spleen, Kidney, Brain and Heart. Bar indicates the estimated molecular weight. The probes with which the same filter was hybridized are indicated in the left side of the panel. (B) Reverse transcriptase/Polymerase chain reaction from different days of embryoid bodies (EB) development. Days are indicated in the upper panel, ES, Stem cells, Fb, fibroblasts. (C) Northern blot using 30µg of total RNA from a pool of embryonic hearts ,E, or adult mouse heart , A, hybridized with the full length mSAP49 probe. Ethidium bromide (EtBr) staining is shown to control for loading and RNA integrity. Note that panel A contains isolated polyA⁺ while panel C shows hybridization signal using total RNA.

7.5-8.0 pc (figure 4B), while the atrial MLC-2a message, used as a control for RNA integrity, was seen in the fusing cardiac tubes (figure 4A). At day 10.0 pc, mSAP49 mRNA was first detected in the optic eminence and optic vesicle (figure 4F). Ridges on the inside of the 4th ventricle (figure 4G) and a weaker signal in the hindbrain and somites were also found (figure 4C, D). At day 11.0 pc, limb and somites appeared strongly stained (figure 4E, F), but no mSAP49 message was detected in the heart or in for -and mid brain structures (figure 4E). Longer incubation of parallel embryo samples confirmed mSAP49 expression in the somites and the negative mSAP49 staining in the heart and brain (figure 4C, D). mSAP49 was first detected in the heart at day 11.5 pc. At this stage it was restricted to the ventricular myocardium. At day 14 pc hybridization signal was also detected

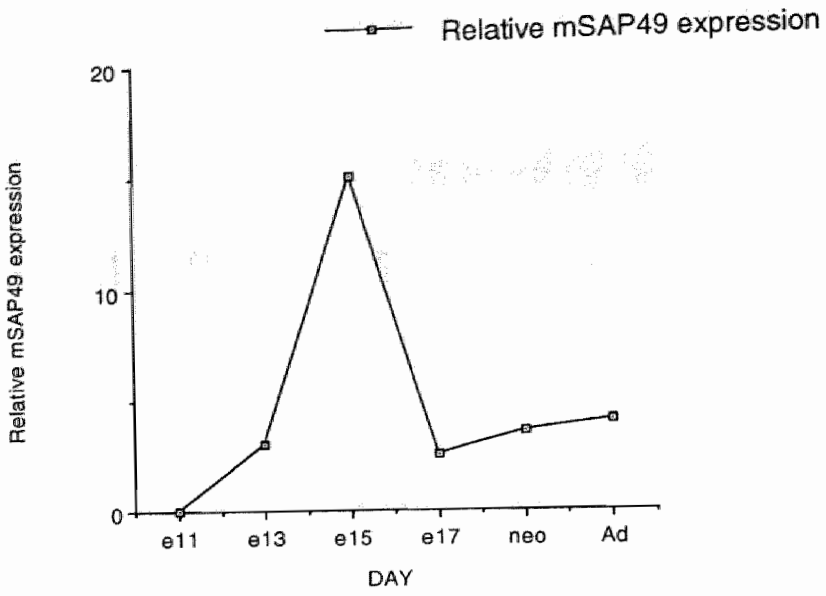


Figure 5. Kinetics of mSAP49 expression in the heart during development. Densitometry analysis of mSAP49 mRNA abundance in isolated hearts is plotted. Measures were taken at embryonic days 11.0 pc, 13.0 pc and 17.0 pc., neonatal and adult mouse hearts. Data were normalized using a GAPDH control probe.

in the atria (figure 4H). The above results suggest a dynamic pattern of mSAP49 expression during cardiogenesis. To complement the *in situ* analysis and quantitatively determine the kinetics of mSAP49 mRNA abundance in heart development, we performed Northern-blot analysis using 30μg of RNA from embryonic heart day 11.0 to the adult cardiac tissue (figure 5). After normalization of the signal by the use of a GAPDH control probe, we determined that mSAP49 displayed a peak of expression at embryonic day 14-15.0 pc and its expression decayed dramatically at day 17.0 pc to reach levels in the neonate comparable to the adult stage.

Human SAP49 was initially cloned as a spliceosome-associated factor although its role in splicing is not completely understood. An essential or general factor of the spliceosome would be expected to be expressed ubiquitously during mammalian development and in a housekeeping manner in the adult. The pattern of mSAP49 mRNA expression during cardiac chamber development suggests that mSAP49 expression is not required in the processing of early cardiac messages. Nevertheless, the transient increase in the mRNA abundance of this pre-spliceosome component coincides with the time frame in which a battery of cardiac proteins change from the embryonic to the adult isoforms[5,25] and it should be considered in this context. Contractile protein transcripts are in the cardiac tube

at the onset of pumping function. At day 8.5 pc of the mouse embryo, mRNAs for the sarcomeric proteins myosin heavy chain α and β, myosin light chain 1a and 1v, cardiac α-actin, tropomodulin, tropomyosin, α-actinin, titin and desmin and the metabolic B-creatine kinase are already detected in the myocardium.[26-28] Nevertheless, the transition between this early pump to the mature cardiac system involves the expression of different myofibrillar, membrane, and cytosolic proteins isoforms than those observed in the embryo. This developmental switch is mediated both by differential splicing of single primary transcripts and transcriptional induction of the adult isoforms.

At day 12.5 pc in the mouse the muscle-specific creatine kinase (MCK) transcript is expressed in the myocardium.[29] Cardiac Troponin I (cTnI) transcripts are detectable in the heart at embryonic day 11 pc in the rat embryo but the protein product is not detected until day 18 pc, suggesting a post-transcriptional mechanism involved in the regulation of (cTnI) expression.[30] Troponin T (TnT) undergoes a perinatal switch from the larger acidic embryonic form to the smaller, basic adult form,[31] which involves differential splicing. Dystrophin gene transcripts reach peak levels around mouse embryonic day 15 pc.[32] These changes have important functional and pathological implications. For example, a fetal TnT isoform is increased in human heart failure, suggesting a decrease in ATPase activity in the failing heart.[33]

Conclusions and implications

We provide the first evidence to date that a spliceosome component (mSAP49) can be developmentally regulated in both a spatial and temporal fashion. The onset of mSAP49 expression in the heart is initially restricted to the ventricle and is seen later in the atria, again suggesting spatial specificity. The fact that this regulation coincides temporally and spatially with the onset of differential expression of a wide variety of alternatively processed transcripts during cardiac development may be more coincidental. A direct test of this notion would require transgenic approaches to express mSAP49 early in the heart tube under the control of a promoter that will drive expression of mSAP49 early in the heart tube, e.g., MLC-2v.

The determination of mSAP49 binding sites *in vitro* using RNA-binding assays[34] along with a systematic determination of the developmental expression of the isolated spliceosomal components will shed important information on the molecular mechanisms that regulate tissue maturation.

Literature

1. Cooper TA, Ohrdahl CP: A single cardiac Troponin T gene generates embryonic and adult isoforms via developmentally regulated alternate splicing. *J Biol Chem* 1985;260:11140-11148
2. Jin JP, Lin JJC: Isolation and characterization of cDNA clones encoding embryonic and adult isoforms of rat cardiac troponin T. *J Biol Chem* 1989;264:14471-14477
3. Breitbart RE, Nadal Ginard B: Developmentally induced, muscle-specific trans factors control the differential splicing of alternative and constitutive Troponin T exons. *Cell* 1987;49:793-803
4. Mullen MP, Smith WJ, Patton JG, Nadal-Ginard B: a-Tropomyosin mutually exclusive exon selection: competition between branchpoint/polypyrimidine tracts determines the default exon choice. *Genes & Development* 1991;5:642-655
5. Schiaffino S, Gorza L, Ausoni S: Troponin isoform switching in the developing heart and its functional consequences. *Trends Cardiovasc Med* 1993;3:12-18
6. Moore MJ, Query CC, Sharp PA: Splicing of precursors to mRNAs by the spliceosome, in Gestland RF, Atkins JF (eds): *The RNA world*. Plainview NY, Cold Spring Harbor Laboratory Press, 1993, pp 303-357
7. Caceres JF, Stamm S, Helfman DM, Krainer AR: Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. *Science* 1994;265:1706-1709
8. Mayeda A, Krainer AR: Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell* 1992;68:365-375
9. Valcarcel J, Singh R, Zamore PD, Green MR: The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. *Nature* 1993;362:171-175
10. Lee KJ, Hickey R, Zhu H, Chien KR: Positive regulatory elements (HF-1a and HF-1b) and a novel negative regulatory element (HF-3) mediate ventricular muscle-specific expression of myosin light-chain 2-luciferase fusion genes in transgenic mice. *Mol Cell Biol* 1994;14:1220-1229
11. Ross RS, Navankasattusas S, Harvey RP, Chien KR: An HF-1a/HF-1b/MEF-2 combinatorial element confers cardiac ventricular specificity and established an anterior-posterior gradient of expression. *Development* 1996;122:1799-1809
12. Zhu H, Nguyen VT, Brown AB, Pourhosseini A, Garcia AV, van Bilsen M, Chien KR: A novel, tissue-restricted zinc finger protein (HF-1b) binds to the cardiac regulatory element (HF-1b/MEF-2) in the rat myosin light-chain 2 gene. *Mol Cell Biol* 1993;13:4432-4444
13. Zou Y, Chien KR: EFIA/YB-1 is a component of cardiac HF-1A binding activity and positively regulates transcription of the myosin light-chain 2v gene. *Mol Cell Biol* 1995;15:2972-2982

14. Miller-Hance WC, LaCorbiere M, Fuller SJ, Evans SM, Lyons G, Schmidt C, Robbins J, Chien KR: In vitro chamber specification during embryonic stem cell cardiogenesis. Expression of the ventricular myosin light chain-2 gene is independent of heart tube formation. *J Biol Chem* 1993;268:25244-25252
15. Christiansen JH, Dennis CL, Wicking CA, Monkley SJ, Wilkinson DG, Wainwright BJ: Murine Wnt-11 and Wnt-12 have temporally and spatially restricted expression patterns during embryonic development. *Mech Dev* 1995;51:341-350
16. Champion Arnaud P, Reed R: The prespliceosome components SAP 49 and SAP 145 interact in a complex implicated in tethering U2 snRNP to the branch site. *Genes Dev* 1994;8:1974-1983
17. Dreyfuss G, Matunis MJ, Pinol Roma S, Burd CG: hnRNP proteins and the biogenesis of mRNA. *Annu Rev Biochem* 1993;62:289-321
18. DeAngelo DJ, DeFalco J, Rybacki L, Childs G: The embryonic enhancer-binding protein SSAP contains a novel DNA-binding domain which has homology to several RNA-binding proteins. *Mol Cell Biol* 1995;15:1254-1264
19. Tay N, Chan SH, Ren EC: Identification and cloning of a novel heterogeneous nuclear ribonucleoprotein C-like protein that functions as a transcriptional activator of the hepatitis B virus enhancer II. *J Virol* 1992;66:6841-6848
20. Grange T, de Sa CM, Oddos J, Pictet R: Human mRNA polyadenylate binding protein: evolutionary conservation of a nucleic acid binding motif. *Nucleic Acids Res* 1987;15:4771-4787
21. King PH, Levine TD, Fremeau RT, Jr., Keene JD: Mammalian homologs of Drosophila ELAV localized to a neuronal subset can bind in vitro to the 3' UTR of mRNA encoding the Id transcriptional repressor. *J Neurosci* 1994;14:1943-1952
22. Feng DF, Doolitle RF: Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J Mol Evol* 1987;25:351-360
23. Rivera-Pomar R, Niessing D, Schmidt-Ott R, Gehring WJ, Jackle H: RNA binding and translational suppression by bicoid. *Nature* 1996;379:746-749
24. Lee KJ, Ross RS, Rockman HA, Harris AN, O'Brien TX, van Bilsen M, Shubeita HE, Kandolf R, Brem G, Price J, et al: Myosin light chain-2 luciferase transgenic mice reveal distinct regulatory programs for cardiac and skeletal muscle-specific expression of a single contractile protein gene. *J Biol Chem* 1992;267:15875-15885
25. Lyons GE: In situ analysis of the cardiac muscle gene program during embryogenesis. *Trends Cardiovasc* 1994;4:70-77
26. Schaart G, Viebahn C, Langmann W, Ramaekers F: Desmin and titin expression in early post-implantation mouse embryos. *Development* 1989;107:585-596
27. Jockusch H, Muller U, Jockusch B: Accumulation and spatial distribution of structural proteins in developing mammalian muscle. *Exp Biol Med* 1984;9:121-125
28. Lyons GE, Schiaffino S, Sassoon D, Barton P, Buckingham ME: Developmental regulation of myosin gene expression in mouse cardiac cells. *J Cell Biol* 1990;111:2427-2436

29. Lyons GE, Muhlebach S, Moser A, Masood R, Paterson B, Buckingham ME, Perriard JC: Developmental regulation of creatine kinase gene expression by myogenic factors in mouse and chick embryos. *Development* 1991;113:1017-1029
30. Gorza L, Ausoni S, Merciai N, Hastings KE, Schiaffino S: Regional differences in troponin I isoform switching during rat heart development. *Dev Biol* 1993;156:253-264
31. Jin JP, Lin JJC: Rapid purification of mammalian cardiac troponin T and its isoform switching in rat hearts during development. *J Biol Chem* 1988;263:7309-7315
32. Houzelstein D, Lyons GE, Chamberlain J, Buckingham ME: Localization of dystrophin gene transcripts during mouse embryogenesis. *J Cell Biol* 1992;119:811-821
33. Anderson PA, Malouf NN, Oakeley AE, Pagani ED, Allen PD: Troponin T isoform expression in the normal and failing human left ventricle: a correlation with myofibrillar ATPase activity. *Basic Res Cardiol* 1992;87 Suppl 1:117-127
34. Siomi H, Siomi MC, Nussbaum RL, Dreyfuss G: The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein. *Cell* 1993;74:291-298

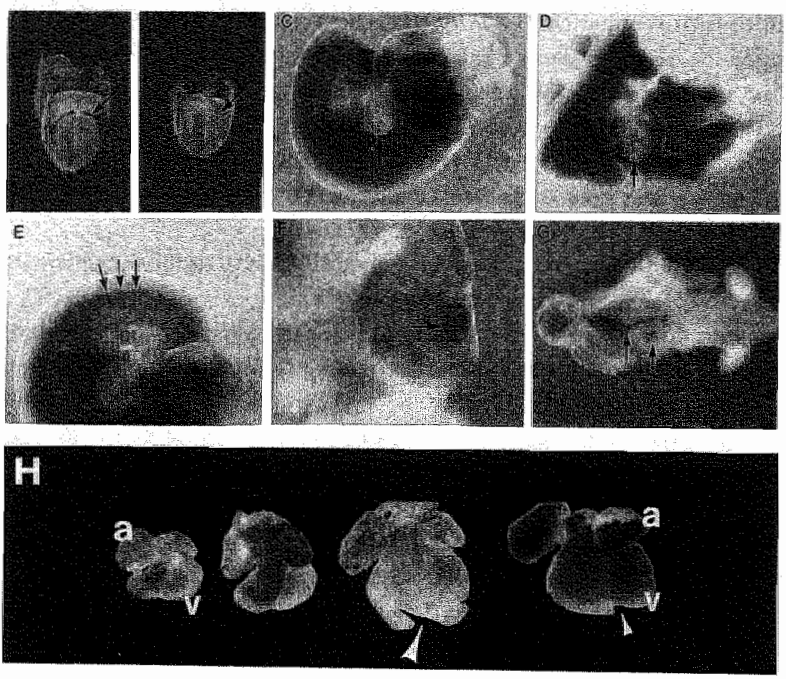


Figure 4. Whole mount *in situ* hybridization in day 8.0 pc (A,B) and day 10.0 pc (C-G) mouse embryos, hybridized with (A) an MLC-2A probe, to control for RNA integrity or (B-G) mSAP49 . Arrows indicate: hybridization signal in somites (E), optic eminence (F) and ridges in the forth ventricle (G). No detectable hybridization signal was observed in the heart (arrows C and D). (H) Whole mount *in situ* hybridization on isolated embryonic hearts. Left to right: e11.5pc, e12.5pc, e13.5pc and e14pc. Symbols indicate: a, atrium; v, ventricle; arrow, small incision in the tissue.

Chapter

Characterization of embryonic stem cell derived cardiomyocytes: a comparative study of *in vivo* and *in vitro* developed murine cardiomyocytes

Pieter A. Doevendans, Rui-Hai An, Steven W Kubalak, Michael Davies, Monique Lacorbiere, Wanda C. Miller-Hance, Kenneth R Chien, Robert S. Kass

Abstract

Background: Embryonic stem cells (D3) will differentiate into embryoid bodies (EB) in the absence of leukemia inhibitory factor. The EB start spontaneous contractions after 4-5 days of differentiation. From these EB, myocytes can be isolated that show several characteristics of cardiomyocytes. This study was performed to further validate the EB derived cardiomyocytes as a potential source of cardiomyocytes for experimental studies. Hereto we studied the expression of the myosin light chain 2 isoforms, as molecular marker of cardiac chamber specification, in relation to the presence of functional ion channels.

Methods: EB were cultured in suspension until a maximum of 16 days after the start of differentiation. Cardiomyocytes were dissociated by collagenase treatment and purified by a density gradient. Standard patch clamp procedures were followed to register cation currents. In addition the cardiomyocytes were studied by immunocytochemistry to study the expression of contractile proteins. In a parallel experiment ion currents were recorded in fetal ventricular cardiomyocytes.

Results: Similar frequency distributions and current characteristics were recorded from EB derived cardiomyocytes, compared to fetal murine cardiomyocytes. Early stage EB-derived cardiomyocytes express the ventricular myosin light chain 2 protein (32%), demonstrated by immunocytochemistry and there is an increase in the number of positive staining cardiomyocytes in late stage EB (44%). There were no chamber specific ion currents recorded in the EB derived cardiomyocytes. Successfully Single cell analysis was used to establish the contractile phenotype of cells after cellular electrophysiology.

Conclusion: The onset of functional ion channel expression in EB derived cardiomyocytes and *in vivo* is comparable. Furthermore the distribution of the ion currents were similar in EB and fetal cardiomyocytes. The EB derived cardiomyocytes express ventricular chamber specific molecular markers.

Introduction

Progress in molecular cardiology is hampered by the lack of cardiomyocyte cell lines. As alternative source for atrial cardiomyocytes AT1 and AT2 cells are available, with stable expression of the simian virus 40 large-T oncogene. H9c2 cells were derived from ventricular embryonic rat heart, and have been described as a cardiomyoblast cell line.[1,2] However this cell line has several features reminiscent of skeletal myocytes. Embryonic carcinoma cells (P19) in the presence of dimethyl sulphoxide (DMSO) form rhythmically contracting muscle cells.[3] Although the H9c2 and P19 cell lines have a cardiomyocyte-like phenotype, these cells do not progress into terminally differentiated atrial or ventricular cardiomyocytes *in vitro*. To circumvent this problem rat neonatal cardiomyocytes can be isolated and cultured to study gene expression in culture conditions.[4-6] In cultured cardiomyocytes gene expression can be modified by pharmacological treatment and transient DNA transfection. For each study a new batch of neonatal rats has to be sacrificed.[6] Furthermore the isolated myocytes exhibit a hypertrophic phenotype, with re-expression of a fetal gene program in culture conditions.[7] A source of differentiating cardiomyocytes could be provided by myocytes isolated from embryoid bodies (EB). These embryoid bodies are formed, when embryonic stem cells in culture are allowed to aggregate and dif-ferentiate.[8] Cardiomyocytes can be isolated from EB by cell dispersion and discontinuous density gradient centrifugation.[9]

This *in vitro* system has many advantages. The undifferentiated stem cells can be cultured at any desired moment and grown in large quantities. Furthermore the ES cells can be manipulated genetically in order to make transgenic or knock-out cell lines. Cardiomyocytes derived from these manipulated cell lines enable investigators to study the functional effect of the genetic intervention, without generating transgenic and gene-targeted animal lines. If an interesting phenotype emerges, the generation of transgenic mouse lines from these modified ES cell lines is relatively easy.[10] A major disadvantage of the system is that potentially any cell type can develop at any site within the EB. Therefore the purity of cells isolated from the EB system is a continuous concern. The atrial and ventricular regulatory myosin light chain (MLC-2) isoforms have been shown to be useful molecular markers for cardiac chamber formation. The ventricular MLC-2(v), is expressed exclusively in ventricular cardiomyocytes, whereas the atrial isoform (MLC-2a) is initially active in all cardiomyocytes but becomes restricted to the atrial chamber at day 12 of murine embryonic development.[11,12]

In this study we validated the EB system as a source of cardiomyocytes using contractile protein expression and ion channel function as potential cardiac chamber specific markers. The results obtained *in vitro* were compared to previously reported *in vivo* results obtained in cultured mouse fetal cardiomyocytes.[13,14] We performed single cell analysis by cellular electrophysiology follow-

ed by immunocytochemistry in this diverse cell population to check for contractile protein isoform expression.

Materials and methods

EB culture and cell purification

The method to harvest cardiomyocytes from EB has been described in detail by Kubalak and coworkers.[15] Briefly, the D3 embryonic stem cell line[8] is maintained and propagated in high glucose Dulbecco's modified Eagle medium (GibcoBRL, Gaithersburg MD) supplemented with 15% heat inactivated fetal calf serum (Sigma Chemical Co, St. Louis, MO), 2 mM L-glutamine and 0.1 mM 2-mercaptoethanol (Sigma). The stem cells are kept in the undifferentiated state by the addition of Leukemia Inhibitory Factor (LIF; ESGRO, Gibco) at 1000 units/ml. The stem cells are allowed to grow in monolayers to near confluency (80%). Cells are maintained at 37°C in a humidified atmosphere of 10% Co_2/ 90% air. *In vitro* differentiation is initiated by harvesting ES cells with trypsin-EDTA and transferring dissociated single cell suspensions to bacterial Petri dishes containing growth medium, as described above, excluding LIF (day 0). Under these conditions cells aggregate to form three dimensional structures the so called EB (embryoid bodies, figure 1), which increase in complexity as differentiation progresses. The EB varying in age from 10 up to day 16 are dispersed by repetitive collagenase treatment. The cardiomyocytes are separated from the mixture of cells using a discontinuous Percoll density gradient. Using a bottom layer of 1.090 g/ml and a top layer of 1.070 g/ml the cardiomyocytes will migrate to the interface of both layers during centrifugation at 3000xg for 30 minutes (figure 2). After removal of the top layer the cardiomyocytes are removed, washed in medium and plated on gridded coverslips (Eppendorf) coated with 20µg/ml laminin and 1% gelatin. For shipment cells were plated in gelatin treated cell culture flasks (65 cm^2) (NUNC, Rosskilde, Danmark) and penicillin/streptomycin was added to the growth medium. After overnight delivery (Fedex, USA) cells were mobilized by trypsin treatment and plated on gridded and fibronectin coated coverslips. Cellular electrophysiology was performed the same day.

Fetal cardiomyocyte culture

A method was developed to culture myocytes derived from day 9.5 to 17 mouse embryos.[9] Embryos were dissected from the uterus, as described by Sturm and Tam.[16] Briefly, pregnant mice are sacrificed by cervical dislocation and the uterus is removed into PBS. The uterus is rinsed once and opened longitudinally to expose the embryos still contained within the decidua. After placing the

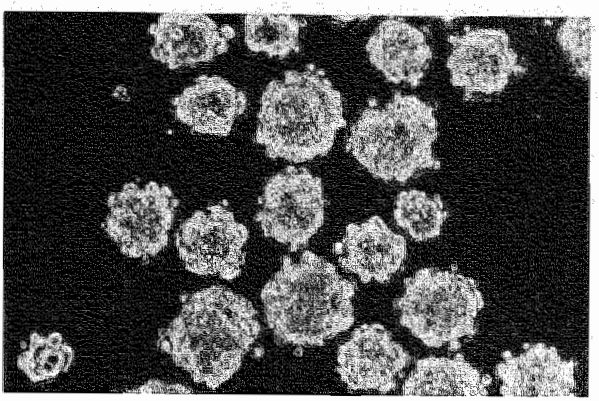


Figure 1A.

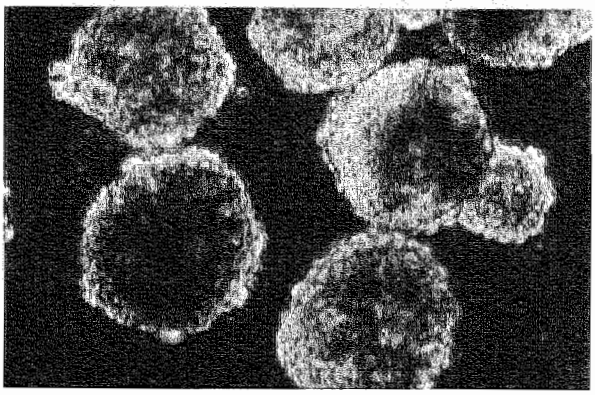


Figure 1B.

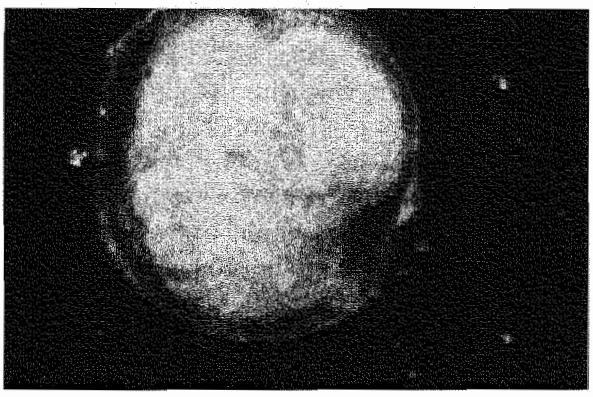


Figure 1C. Embryoid body (EB) formation. Note the transparent aggregate of embryonic stem cells, forming multicellular bodies. These structures can start spontaneous contractions after 4-5 days of differentiation. Three different stages of EB formation are shown A) Day 5, size 0.2-0.3 mm B) Day 10, size 0.5-0.7 mm. C) Day 16, size 40mm (magnification approximately 10x). Note the cyste formation in the later stage EB.

embryos in ADS buffer the hearts are removed from the embryo. The ADS buffer is made with 116 mM NaCl, 20 mM N-2-hydroxethylpiperazine-N'-2-ethanesulfonic acid (HEPES), 1.0 mM NaH_2PO_4, 5.0 mM KCl, 0.8 mM $MgSO_4$, 5.5 mM glucose, pH 7.35, and filter sterilized with a 0.22-μm pore size filter. The atria can be distinguished by localization and color and separated from the heart by pulling both atria free with a forceps (figure 3). The tissue is placed in a Eppendorf tube spun briefly to replace the solution with fresh ADS buffer containing 0.5 mg/ml Collagenase type II (Worthington Biochemical Corporation, Lakewood NJ) and 1.0 mg/ml pancreatin (GibcoBRL). Digestion is allowed for 15 min. at 37°C on a rotating Ferris wheel followed by gentle trituration of the tissue. Tubes are quick-spun and the supernatant containing isolated myocytes is transferred to DMEM supplemented with 10% horse serum and 5% fetal calf serum. If necessary these steps can be repeated once. The cells were plated on gelatin and laminin treated coverslips (1% gelatin and 20 μg/ml Laminin; Sigma). The majority of cells will exhibit vigorously beating after a few hours and start dividing after 1-2 days in culture.

Patch clamping

For this study the conventional whole cell configuration was used.[17] Both beating and non-beating isolated embryoid body derived cardiomyocytes (EBDC) were used. Intracellular and extracellular solutions, designed to isolate K^+, Ca^{2+}, and Na^+ channel currents have been described previously.[18,19] Patch pipettes (Clay Adams glass) were pulled to resistances of 2.5 to 5.0 MΩ, when filled with intracellular solutions. Capacity transients were monitored as a function of time after attaining a high resistance seal with the surface membrane, and electrical access was usually reached within 10 minutes of seal formation. Cell capacitance was determined either by integration of the capacity transient caused by 5-mV voltage step or analog measurement with the patch clamp amplifier. Distinct pulse protocols were chosen for each channel studied. In general, negative (-80 or -90 mV) holding potentials were used for investigation of Na^+ channel currents, and inactivating K^+ channel currents. Tetrodotoxin (10 μmol/L, Sigma) was added to solutions in which Na^+ channels were to be eliminated, but holding potentials had to be negative. Nisoldipine (200 nmol/L) was added to solutions in experiments where L-type Ca^{2+} channel activity needed to be eliminated. More depolarized holding potentials (-50 or -40 mV) were generally used for measurement of L-type Ca^{2+} channels, delayed K^+ channels, and inwardly rectifying K^+ channels. These more positive holding potentials resulted in the inactivation of rapidly activating/inactivating K^+ channels.

Data were collected, stored, and analyzed on IBM 486-compatible computers interfaced to Yale Mark IV amplifiers, constructed in our laboratory or to an Axopatch 200A amplifier driven by PCLAMP software (Axon Instruments,

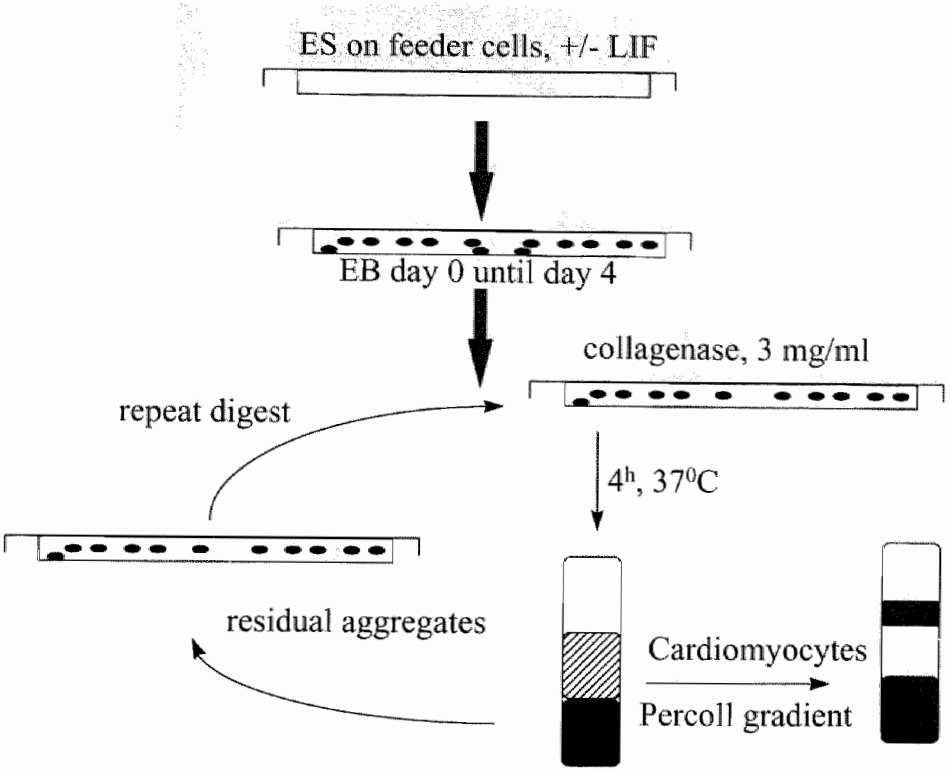


Figure 2. Schematic representation of isolation procedure of embryoid body derived cardiomyocytes. Embryonic stem cells are grown in the presence of leukemia inhibitory factor to near confluency. The cells are mobilized and replated in the absence of LIF and feeder cells. After 11-16 days of differentiation embryoid bodies were harvested, and dispersed by repetitive collagenase treatment. Embryoid body derived cardiomyocytes were purified by a discontinuous Percoll gradient centrifugation.

Burlingame, CA). Graphics and statistical data analysis were carried out by ORIGIN software (Microcal, Northampton, MA).

Immunocytochemistry

Immunocytochemistry was performed to check the contractile protein expression pattern in developing EBDC. Part of the cells isolated for electrophysiologic studies were stained to check the cardiomyocyte phenotype as indicated by the expression of atrial myosin light chain 2 (MLC-2a) and ventricular MLC-2 (MLC-2v). The cells were co-immunostained with α-myomesin to recognize all myocytes. In three different cultures the number of myocytes positive for either MLC-2a or MLC-2v per 100 α-myomesin positive cells were counted.

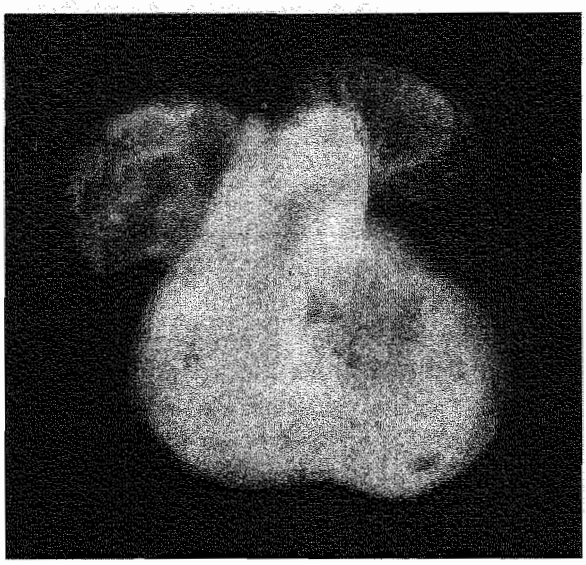


Figure 3. Picture of embryonic murine heart day 12 post coitum. Note the clear demarcation of the atrial appendices at this early developmental stage (actual size approximately 1.5 mm).

The procedure followed to determine the contractile phenotype of the EBDC following an electrophysiologic examination in brief; after termination of the patch clamp experiment cells were fixed by local perfusion with 4% paraformaldehyde for 1 min. After this initial fixation the pipet could be withdrawn, leaving an apparently normal shaped EBDC on the gridded coverslip. The method followed thereafter was identical to previously described methods of immunostaining.[9] Cells were washed in PBS (2x) and fixed in 4% paraformaldehyde for 5 min. The solution was neutralized with 50 m*M* NH_4Cl, and the cells were permeabilized with 0.1% Triton X-100 in PBS (PBST) with 1% bovine serum albumin for 30 min. The primary antibody solution contained PBST with a mixture of murine monoclonal anti-α-myomesin (1:200 dilution) and either anti-atrial MLC-2 (1:500) or anti-ventricular MLC-2 (1:60). The MLC-2a and MLC-2v antibodies are both polyclonal rabbit antibodies. After 4 washes in PBS the cells were blocked during 30 min. with 1% BSA and 1% donkey serum in PBST. The secondary antibodies FITC donkey-anti-rabbit and Texas Red goat-anti-mouse were applied for 1 hour. The cells were washed again 4 times and mounted with Gelvatol supplemented with triethylenediamine. The whole procedure was performed at room temperature (22°C). The application of the secondary antibodies and subsequent washes were performed in the dark to protect the im-

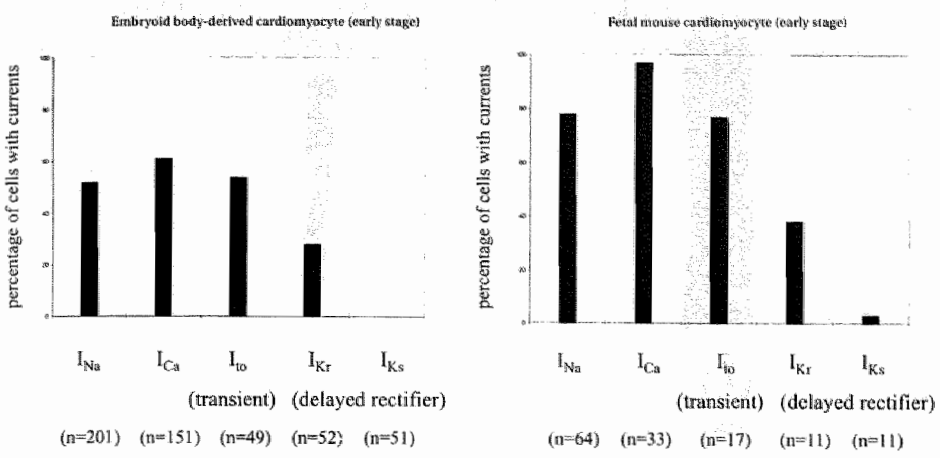


Figure 4. Distribution of ionic currents in embryoid body derived cardiomyocytes and fetal ventricular cardiomyocytes. Bar diagram showing the percentage of positive cells (Y-axis) in relation to the total number of cells(n) studied for the specific current (X-axis). The left panel shows the embryoid body derived cardiomyocytes, the right represents the results from the fetal cardiomyocytes. Note the low percentage of cells in which I_{Ks} was recorded.

munofluorescent signal. For microscopy, standard epifluorescent optics were used to study the slides the next day.

Results

EBDC were obtained at day 11-12 (early) and 15-16 (late) of embryoid body formation after the change to growth medium. The process of differentiation starts when the ES cells are grown without leukemia inhibitory factor and in the absence of a fibroblast feeder layer. The expression and regulation of ion channel currents in EBDC and fetal murine ventricular cells was compared, focusing on L-type calcium (Ca^{2+}), voltage-gated sodium, and inactivating potassium channel currents.

Distribution of currents in EBDC derived cardiomyocytes

As described previously for early stage (day 11-12) fetal ventricular cells, we found that EBDC express both voltage-gated Na^+ channel and L-type Ca^{2+} channel currents. The percentage of EBDC showing the excitatory currents ($Ca^{2+}$and Na^+) relative to the number of total cells tested is shown in figure 4. Approximately 50% of the EBDC tested had Na^+ currents, while more often (±

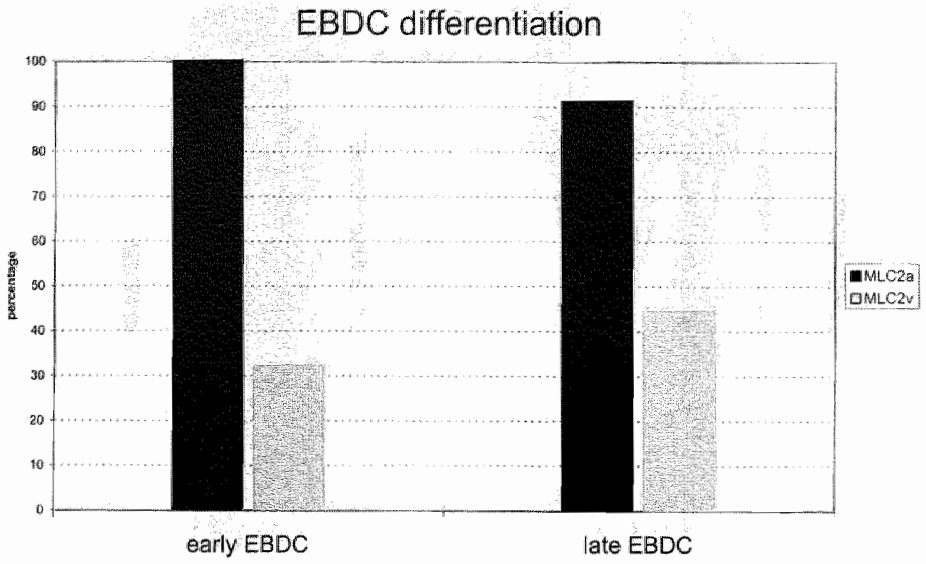


Figure 6. The frequency of expression of myosin light chain-2a (MLC-2a) and myosin light chain 2v (MLC-2v) protein in early (day 11) and late (day 16) embryoid body derived cardiomyocytes. In three independent cultures 100 α-myomesin positive cells were scored co-stained with either MLC-2a or MLC-2v antibodies. The bar graphs show the percentage of positive cells for both MLC-2 antibodies at the two stages.

Table 1. Currents in relation to contractile protein expression in single cells, isolated and purified from EB.

	Antibodies	Currents	Number of cells analyzed
I)	Myomesin -	Absent	32
II)	Myomesin + MLC-2a -	–	0
	MLC-2a + (16)	I_{Ca},I_{Na},	10
		I_{to},I_{Kr}	6
III)	Myomesin + MLC-2v - (10)	I_{Ca},I_{Na},	8
		I_{to},I_{Kr}	2
	MLC-2v + (8)	I_{Ca},I_{Na},	5
		I_{to},I_{Kr}	3

Embryoid body derived cardiomyocytes were evaluated for the presence of either I_{Ca},I_{Na}, or I_{to},I_{Kr}. Classification: Myomesin -: non-myocyte (I); Myomesin +, MLC-2a + or MLC-2v -; undifferentiated or atrial myocyte (II, III). Myomesin +, MLC-2v +: ventricular myocyte (III). It is noteworthy that all examined myomesin positive cells were positive with the MLC-2a antibody when present. For myomesin detection we used the murine monoclonal α-myomesin antibody. The myosin light chain 2 antibodies are polyclonal rabbit-anti-mouse antibodies.

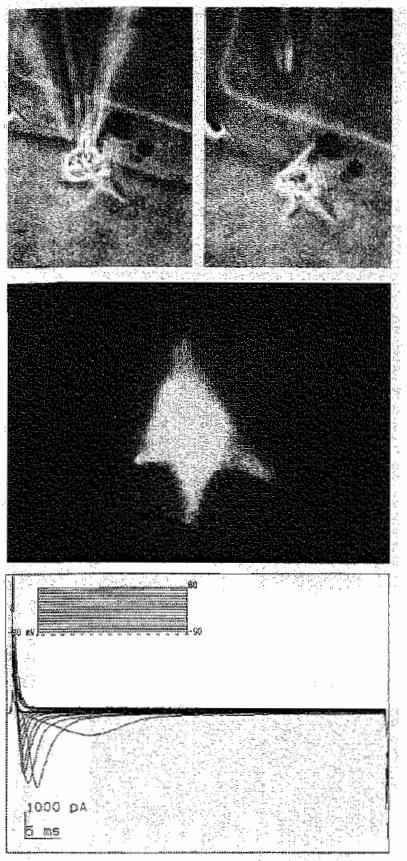


Figure 7. Single cell analysis: EBDC during (A) and after electrode withdrawal (B, magnification 10x). The same cell after immuno-staining (C). Shown is a digital image of MLC-2a antibody stained with FITC labeled goat-anti-rabbit as the secondary antibody (magnification 40x). (D) The current trace of I_{Na} derived from this embryoid body derived cardiomyocyte is shown. Presented is a family of current tracings induced by a series of depolarizing pulses from a holding potential of -80 mV to test pulses between -80 and +40 mV in a day 12 embryoid body derived cardiomyocyte.

60%) L-type Ca^{2+} currents were recorded. In early stages of the fetal mouse cardiomyocytes close to 80% of the cells showed Na^+ and 100% had characteristic Ca^{2+} currents. In both systems therefore some myocytes show Ca^{2+} currents in the absence of recordable Na^+ currents.

The distribution of the repolarizing currents is shown in figure 4 as well. Again the frequency distribution of I_{to}, I_{Kr} and I_{Ks} currents is remarkable similar in both systems. The lack of I_{Ks} currents is a consistent finding in both types of cardiomyocytes. The difference in the number of cells that have currents when comparing EBDC to fetal cardiomyocytes is due to the difference in cardiomyocyte isolation. The purification of cardiomyocytes from EB through the Percoll gradient resulted in cultures containing 50-60% cardiomyocytes as assessed by immunocytochemistry. In early and late EBDC, identical currents were recorded, showing identical frequency distribution and current characteristics (data not shown).

Figure 5A.

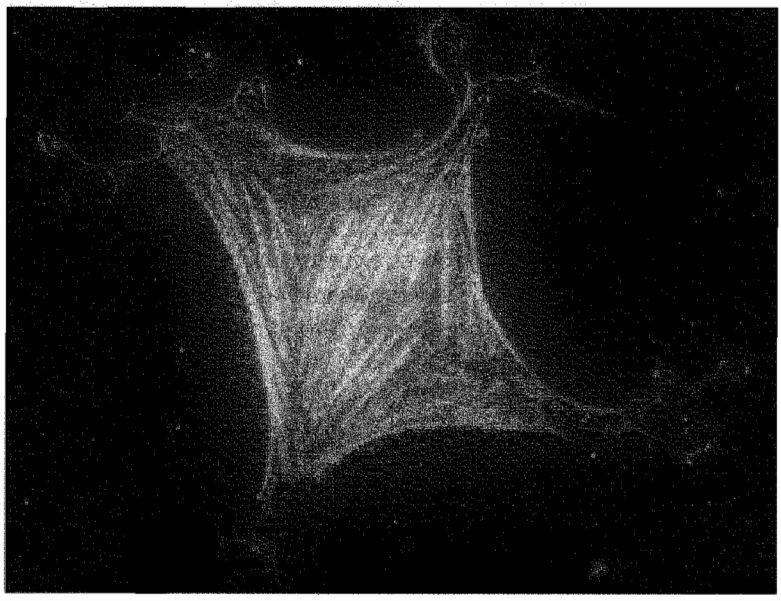


Figure 5B. Immunostaining of isolated embryoid body derived cardiomyocyte (day 11) with the polyclonal atrial myosin light chain 2 antibody (panel A) and the polyclonal ventricular myosin light chain 2 antibody (panel B, day 16, magnification 100x). The secondary antibody was fluorescein isothiocyanate labeled goat-anti-rabbit immunoglobin.

Immunocytochemistry

α-Myomesin positive cells staining with MLC 2v (figure 5A) and cells negative for MLC-2a were considered ventricular myocytes. Cells positive for MLC 2a (figure 5B) or negative for MLC 2v can be considered to be either atrial myocytes, or undifferentiated cardiomyocytes. Approximately 50-60% of all the purified cells were positive for α-myomesin. The distribution of each cell type in three different cultures is shown in figure 6. Approximately 32% of the early EBDC had the ventricular phenotype versus 44% in the late EBDC.

Single cell analysis

The control experiment to check for contractile protein expression in relation to current recordings in single cells, revealed identical currents in all cardiomyocytes studied (table 1). A total of 64 cells was included for this analysis. No currents were registered in half (32) of the cells. Only cells that showed cross striations or a filamentous staining pattern with the monoclonal α-myomesin, were considered cardiomyocytes. Among the cells isolated and purified from EB, in which no currents were recorded, neither α-myomesin nor MLC-2 positive cells were found. The immunocytochemistry was suggestive for an undifferentiated or atrial phenotype in 24 of the cells with currents. In these cells we recorded I_{Ca},I_{Na},I_{to}, and I_{Kr}. No atrial specific current like I_{KACh} was documented in EBDC. Eight cells stained positive for both α-myomesin and MLC-2v and therefore were considered to be ventricular cardiomyocytes, and identical currents (I_{Ca},I_{Na},I_{to}, and I_{kr}) were recorded. No difference was found in current characteristics comparing atrial\non differentiated cardiomyocytes with ventricular cardiomyocytes. An example of a control experiment is shown in figure 7. The current characteristics for the Na^+ (panel D) and Ca^{2+} (not shown) currents are recorded in a cell staining positive for both α-myomesin (not shown) and atrial MLC-2 (panel C).

Discussion

Contractile protein expression in EBDC

The pattern of contractile protein expression in developing embryonic hearts has been described extensively for rodent species. Fetal myocytes express the β-cardiac myosin heavy chain, which after birth is partially replaced by the α isoform.[20] A similar pattern of expression has been reported for the troponin I gene. Initially the slow skeletal muscle isoform is detected in the heart, followed by expression of the cardiac isoform after birth.[21,22] For tropomyosin the situation

is complex as 4 different genes and alternative splicing provide a range of different proteins. Two tropomyosin isoforms (α and β) are present in the heart from early stages and remain active throughout development into the adult rodent heart.[23] The situation is different for the myosin light chain-2 isoforms. In all developmental stages the MLC-2v gene is expressed exclusively in ventricular cardiomyocytes.[12] The atrial isoform MLC-2a is initially found in the entire heart tube but becomes restricted to the atrial chamber shortly after mid gestation (day 12).[11] The expression pattern studies of the contractile proteins in embryoid bodies have shown an identical pattern during ongoing differentiation. Both the α-myosin heavy chain[8] and cardiac troponin I[24] are found in differentiated EB, 20 days after start of beating. The expression pattern of the tropomyosin gene is identical *in vivo* and *in vitro*.[23]

Both the atrial and ventricular isoform of myosin light chain-2 gene are expressed in EB.[11,25] In this report it is shown that early EBDC are 100% positive for the MLC-2a protein and only a limited number of cells (± 30%) have the ventricular isoform. During progressive EB differentiation the percentage of MLC-2v positive cells increases relative to the MLC-2a expressing cells. Again this pattern is reminiscent of the *in vivo* situation. Although *in vivo* the percentage of ventricular myocytes exceeds the number of atrial myocytes at later developmental stages due to organ formation. The absence of organ formation may cause a temporal delay in the differentiation of cardiomyocytes. This appears to be contradictive with the finding of contractile protein isoforms like α-myosin heavy chain and cardiac troponin I, normally expressed after birth *in vivo*. Clearly the potential for full differentiation into ventricular cardiomyocytes is present in this *in vitro* system, but positional clues may be crucial for the recruitment of cells and timing of differentiation into the ventricular phenotype. Support for this theory came from the work by Chen et al.[26] They implanted developing EB in adult mice at different sites and showed a relation between the site of implantation and the final cellular composition of the growing and differentiating EB.

α-Myomesin is present in the M-band of skeletal muscle and heart. α-Myomesin is believed to anchor titin to myosin in the thick filaments. The developmental expression of the murine gene has not been described up to now. In chicken a skeletal and cardiac specific isoform derived from one gene through alternative splicing were recently described. The earliest onset of α-myomesin expression in the chicken heart is currently being evaluated.[27] The mRNA for the murine MLC-2a gene has been detected as early as day 8 post coitum. In this study we found no MLC-2a immunoreactivity in α-myomesin negative cells, suggesting that both α-myomesin and MLC-2a are expressed early in cardiogenesis. Whether α-myomesin protein precedes the expression of functional ion channel proteins was not addressed in this study.

Functional studies in EBDC

Recently several groups started functional assays to study EBDC proteins. Both cellular electrophysiology and measurements of contractility were applied.[28,29] For these studies beating areas were dissected from EB and used for cell isolation. Using the method of beating cell selection EBDC can be maintained in culture for more than 100 days, after the initiation of spontaneous contractions. In a report from Maltsev et al.[30] the isolated EBDC were classified as early, intermediate and late depending on the onset of contractile activity. The late EBDC reported here would match the intermediate stage of development described by Maltsev et al. The characteristics and distribution of the excitatory and repolarizing currents registered from intermediate stage EBDC, are comparable to our results. In addition they reported the presence of I_{KACh},$I_{K,ATP}$,and I_f. These channels were found in cells isolated between 16 and 25 days (terminal differentiation stage[30]) after start of differentiation and were used to classify cells as being atrial, ventricular or nodal in phenotype respectively. In our studies of EBDC we did not find I_{KACh},$I_{K,ATP}$,and I_f. There are two possible explanations for this difference in results. First of all the latest EBDC studied in this report was day 16 after initiation of differentiation (independent of contraction initiation). The presence of the I_{KACh}, in atrial cardiomyocytes *in vivo* is found after day 17.[14] Secondly, the method of cell selection applied is essentially different. The gradient purification step did not result in a pure cardiomyocyte preparation. It is likely that for instance nodal cells may be lost during the purification process. The advantage of the EB suspension system, together with the Percoll purification method is the possibility to isolate a large number of cardiomyocytes in one experiment within a reasonable time frame.

Different studies have suggested a relation between the initiation of spontaneous contraction and changes in protein expression.[29,30] In our experience the percentage of beating EB is less important for differentiation then the number of days in culture. The yield of cardiomyocytes expressing the ventricular phenotype is similar from separate cultures and not dependent on the percentage of beating EB. Up to now it is unclear through which mechanism spontaneous contractions occur. Also the level of differentiation required for spontaneous beating activity has not been established. Therefore the presumption that the EB exhibiting spontaneous contractions must be at the same level of differentiation has not been proven unequivocally.

One of the interesting findings of this study is the absence of I_{Ks}. The potassium currents I_{Kr} and I_{Ks} together are responsible for delayed rectification. We recorded I_{Kr} in about 30% of the EBDC studied, while no I_{Ks}[18] was found in EBDC, and only in a small number of early fetal cells. The absence of the I_{Ks} makes these EBDC and early fetal cardiomyocytes unique substrates to study the effect of

drugs on I_{Kr} activity, but more importantly, these preparations become excellent systems to investigate the factors that control the expression of I_{Ks}.

The contractile function of EBDC was studied by Metzger et al.[29] They compared spontaneous force generation in microdissected EBDC attached to an ultrasensitive force transducer and galvanometer. The transition of cardiac contractile sensitivity to calcium was determined after permeabilization of the cellular membrane. Tension development in cells derived from early stages (beating 6 days) were more sensitive to Ca^{2+} treatment, then later stages (>35 days). These functional measurements are consistent with the results found *in vivo* and once more underline the validity of this *in vitro* source of cardiomyocytes.[29,31] The molecular basis for this altered Ca^{2+} sensitivity is not clear. Switches in isoform expression of the contractile proteins or an increasing number of cells differentiating into the ventricular phenotype might provide an explanation. The decrease in Ca^{2+} sensitivity potentially, could be important in optimizing the cardiac contractile response to variations in the cytoplasmic Ca^{2+} transients.

Single cell analysis

The current analysis is based on the assumption that α-myomesin recognizes only cardiomyocytes in cells isolated from EB and that skeletal myocytes, are not present at the stage examined. This notion is supported by the results in several reports on contractile protein expression in D3 ES cell derived EB.[24,32] Furthermore Miller Hance et al.[25] showed the late onset of myogenic factor expression. MyoD, and myogenin were detected by rtPCR from day 16 of EB development. These genes represent dominant transcription factors required to initiate the skeletal muscle gene programme. Sofar there are no reports on skeletal muscle formation in the D3 stemcell line, but potentially skeletal muscle cells can originate in this system in later stages of development. The α myomesin protein is exclusively localized in the M-band of cross-striated muscle cells.[33-35] In the absence of skeletal muscle development the cells expressing α-myomesin are cardiomyocytes. In this study we found no α-myomesin, or MLC-2 positive cells, among the cells lacking currents. This finding supports the approach to identify EBDC by the expression of α-myomesin.

To our knowledge this is the first report, where immunocytochemistry was applied following cellular electrophysiology to verify the contractile protein expression and, hence, the phenotype of the cardiomyocyte under investigation. A brief fixation (1 min) before electrode withdrawal, preserved cellular morphology. This method was developed to allow further characterization of the contractile apparatus composition of an EBDC in relation to ion channel expression. The establishment of MLC-2v versus MLC-2a isoforms expressed allows an early distinction between differentiated ventricular versus nondifferentiated/atrial cardiomyocytes in this system. At the early stages of EBDC evaluated in this study

we found no difference in ion channel expression in relation to MLC-isoform expression. Alternative methods for single cell analysis are single cell PCR, *in situ* hybridization, or *in situ* PCR[36]. Single cell PCR would allow the simultaneous identification of more then one transcript. Single cell immunocytochemistry, however, determines the presence and localization of the protein of interest.

Applications of ES cells and EBDC

The ES cell system is amenable to genetic manipulation. The generation of transgenic and gene targeted ES cell lines has become a straightforward procedure by now. The extensive validation of the EBDC developing *in vitro* shows that this system can be applied to address questions of early cardiogenesis. In some experiments studies in EBDC could replace the generation of transgenic, knock-out or knock-in mouse lines, thereby reducing time and cost considerably. Furthermore the genetically manipulated ES cells can always be stored and used later to generate mouse lines if required for further *in vivo* analysis. This study clearly shows that for experiments addressing ion channel function important information can be obtained from EBDC.

The transparency of the developing EB and the development of *in vivo* reporter genes allows interesting approaches to visualize developing EB structures. The introduction of *in vivo* reporter genes driven by a cardiac specific promoter like MLC-2v or βMHC would help to resolve matters as the timing of spontaneous contractile function in relation to promoter activation. Furthermore early stages of cardiomyocytes could be purified by cell sorting for gene expression or grafting studies.[37] The potential therapeutical use of *in vitro* developed cardiomyocytes was explored by Klug et al.[38] A stable ES cell line was generated harboring a fusion gene consisting of the α myosin heavy chain promoter driving the cDNA of the aminoglycoside transferase. Expression of this gene will render cells resistant for antibiotic treatment with drug G418. The ES cell line was allowed to differentiate for 8 days, followed by G418 incubation for 8 days. The surviving cells appeared to contain 99.6% cardiomyocytes. These cardiomyocytes were injected into murine ventricular myocardium and produced stable intracardiac grafts. The effect of this procedure on cardiac function remains to be determined.

Conclusions

Our studies support the value of the EB system to study murine cardiogenesis. There were no qualitative differences between cardiomyocytes developing *in vivo* or *in vitro*. In the studied cell population from day 12-16 EBs we found no differences in ion channel expression comparing EBDC with fetal cardiomyocytes that would allow the unequivocal discrimination of cardiomyocytes diffe-

rentiated to the ventricular phenotype. A new method for single cell analysis was developed by preserving cells after electrophysiology for phenotyping with specific antibodies.

Literature

1. Kimes.B, Brandt.B L: Properties of a clonal muscle cell line from rat heart. *Exptl Cell Res* 1976;98:367-381
2. Sipido KR, Marban E: L-type calcium channels, potassium channels, and novel nonspecific cation channels in a clonal muscle cell line derived from embryonic rat ventricle. *Circ Res* 1991;69:1487-1499
3. McBurney MW, Jones-Villeneuve EMV, Edwards EMS, Anderson PJ: Control of muscle and neuronal differentiation in a cultured embryonal carcinoma cell line. *Nature* 1982;299:165-167
4. Shubeita HE, McDonough PM, Harris AN, Knowlton KU, Glembotski CC, Brown JH, Chien KR: Endothelin induction of inositol phospholipid hydrolysis, sarcomere assembly, and cardiac gene expression in ventricular myocytes. A paracrine mechanism for myocardial cell hypertrophy. *J Biol Chem* 1990;265:20555-20562
5. Shubeita HE, Martinson EA, van Bilsen M, Chien KR, Brown JH: Transcriptional activation of the cardiac myosin light chain 2 and atrial natriuretic factor genes by protein kinase C in neonatal rat ventricular myocytes. *Proc Natl Acad Sci U S A* 1992;89:1305-1309
6. Simpson P, Shoshana S: Differentiation of rat myocytes in single cultures with and without proliferating non-myocardial cells. *Circ Res* 1982;50:101-116
7. van Bilsen M, Chien KR: Growth and hypertrophy of the heart: towards an understanding of cardiac specific and inducible gene expression. *Cardiovasc Res* 1993;27:1140-1149
8. Robbins J, Gulick J, Sanchez A, Howles P, Doetschman T: Mouse embryonic stem cells express the cardiac myosin heavy chain genes during development in vitro. *J Biol Chem* 1990;265:11905-11909
9. Kubalak S, Doevendans PA, Rockman H, Hunter JJ, Tanaka N, Ross J, Chien KR: Molecular analysis of cardiac muscle diseases based on mouse genetics, in Adolph KW (ed): *Human molecular genetics*. Orlando, Academic Press, 1996, pp 470-487
10. Dzau VJ, Gibbons GH, Kobilka BK, Lawn RM, Pratt RE: Genetic models of human vascular disease. *Circulation* 1995;91:521-531
11. Kubalak SW, Miller-Hance WC, O'Brien TX, Dyson E, Chien KR: Chamber specification of atrial myosin light chain-2 expression precedes septation during murine cardiogenesis. *J Biol Chem* 1994;269:16961-16970
12. O'Brien TX, Lee KJ, Chien KR: Positional specification of ventricular myosin light chain 2 expression in the primitive murine heart tube. *Proc Natl Acad Sci U S A* 1993;90:5157-5161

13. An RH, Davies MP, Doevendans PA, Kubalak SW, Bangalore R, Chien KR, Kass RS: Developmental changes in beta-adrenergic modulation of L-type Ca2+ channels in embryonic mouse heart. *Circ Res* 1996;78:371-378
14. Davies MP, An RH, Doevendans P, Kubalak S, Chien KR, Kass RS: Developmental changes in ionic channel activity in the embryonic murine heart. *Circ Res* 1996;78:15-25
15. Iannaccone PM, Taborn GU, Garton RL, Caplice MD, Brenin DR: Pluripotent embryonic stem cells from the rat are capable of producing chimeras. *Dev Biol* 1994;163:288-292
16. Shani M, Faerman A, Emerson CP, Pearson White S, Dekel I, Magal Y: The consequences of a constitutive expression of MyoD1 in ES cells and mouse embryos. *Symp Soc Exp Biol* 1992;46:19-36
17. Horn R, Marty A: Muscarinic activation of ionic currents measured by a new whole-cell recording method. *J Gen Physiol* 1988;92:145-159
18. Freeman LC, Kass RS: Delayed rectifier potassium channels in ventricle and sinoatrial node of the guinea pig: molecular and regulatory properties. *Cardiovascular Drugs & Therapy* 1993;7 Suppl 3:627-635
19. Kass RS, Bangalore R: Measurement of dihydropyridine modulation of cardiac L-type calcium channels: molecular and cellular implications, in Narahashi T (ed): *Methods in neurosciences: ion channels of excitable cells*. New York, NY, Academic Press, 1994, pp 189-204
20. Madhavi V, Chambers AP, Nadal Ginard B: Cardiac alpha-and beta- myosin heavy chain genes are organized in tandem. *Proc Natl Acad Sci U S A* 1984;81:2626-2630
21. Bhavsar P, Dhoot G, Cumming D, Butler-Browne G, Yacoub M, Barton P: Developmental expression of the troponin I isoform in the fetal human heart. *FEBS lett* 1991;292:5-8
22. Martin A, Ball K, Gao L, Kumar P, Solaro R: Identification and functional significance of troponin I isoforms in neonatal rat myofibrils. *Circ Res* 1991;69:1244-1252
23. Muthuchamy M, Pajak L, Howles P, Doetschman T, Wieczorek DF: Developmental analysis of tropomyosin gene expression in embryonic stem cells and mouse embryos. *Mol Cell Biol* 1993;13:3311-3323
24. Westfall MV, Samuelson LC, Metzger JM: Troponin I isoform expression is developmentally regulated in differentiating embryonic stem cell-derived cardiac myocytes. *Dev Dyn* 1996;206:24-38
25. Miller Hance WC, LaCorbiere M, Fuller SJ, Evans SM, Lyons G, Schmidt C, Robbins J, Chien KR: In vitro chamber specification during embryonic stem cell cardiogenesis. Expression of the ventricular myosin light chain-2 gene is independent of heart tube formation. *J Biol Chem* 1993;268:25244-25252
26. Chen U, Kosco M: Differentiation of mouse embryonic stem cells in vitro: III. Morphological evaluation of tissues developed after implantation of differentiated mouse embryoid bodies. *Dev Dyn* 1993;197:217-226

27. Bantle S, Keller S, Haussmann I, Auerbach D, Perriard E, Muhlebach S, Perriard JC: Tissue-specific isoforms of chicken myomesin are generated by alternative splicing. *J Biol Chem* 1996;271:19042-19052
28. Maltsev VA, Rohwedel J, Hescheler J, Wobus AM: Embryonic stem cells differentiate in vitro into cardiomyocytes representing sinusnodal, atrial and ventricular cell types. *Mech Dev* 1993;44:41-50
29. Metzger JM, Lin WI, Samuelson LC: Transition in cardiac contractile sensitivity to calcium during the in vitro differentiation of mouse embryonic stem cells. *J Cell Biol* 1994;126:701-711
30. Maltsev VA, Wobus AM, Rohwedel J, Bader M, Hescheler J: Cardiomyocytes differentiated in vitro from embryonic stem cells developmentally express cardiac-specific genes and ionic currents. *Circ Res* 1994;75:233-244
31. Metzger JM, Samuelson LC, Rust EM, Wetsfall MV: Embryonic stem cell cardiogenesis: applications for cardiovascular research. *Trends Cardiovasc Med* 1997;7:63-68
32. Dekel I, Magal Y, Pearson White S, Emerson CP, Shani M: Conditional conversion of ES cells to to skeletal muscle by an exogenous MyoD1 gene. *New Biol* 1997;4:217-224
33. Carlsson E, Grove BK, Wallimann T, Eppenberger HM, Thornell LE: Myofibrillar M-band proteins in rat skeletal muscles during development. *Histochemistry* 1990;95:27-35
34. Perriard JC, von Arx P, Bantle S, Eppenberger HM, Eppenberger Eberhardt M, Messerli M, Soldati T: Molecular analysis of protein sorting during biogenesis of muscle cytoarchitecture. *Symp Soc Exp Biol* 1992;46:219-235
35. Schultheiss T, Lin ZX, Lu MH, Murray J, Fischman DA, Weber K, Masaki T, Imamura M, Holtzer H: Differential distribution of subsets of myofibrillar proteins in cardiac nonstriated and striated myofibrils. *J Cell Biol* 1990;110:1159-1172
36. Surmeier DJ, Wilson CJ, Eberwine J: Patch-clamp techniques for studying potassium currents in mammalian brain neurons, in Conn PM (ed): *Methods in neuroscience.* Academic Press, 1994, pp 39-67
37. Doevendans PA, Becker DA, An R, Kass RS: The utility of fluorescent *in vivo* reporter genes in molecular cardiology. *Biochem Biophys Res Comm* 1996;222:352-358
38. Klug MG, Soonpaa MH, Koh GY, Field LJ: Genetically selected cardiomyocytes from differentiating embryonic stem cells form stable intracardiac grafts. *J Clin Invest* 1996;98:216-224

Chapter

Cloning of the murine atrial myosin light chain-2 gene: a member of an evolutionary conserved family of contractile proteins

Pieter A. Doevendans, Ronald Bronsaer, Ming Dong Zhou, Steve Kubalak, Pilar Ruiz Lozano, Kenneth R Chien, Marc van Bilsen

Abstract

Background: Recently the near complete cDNA of the murine regulatory atrial myosin light chain (MLC-2a) was cloned. Initially the gene is found in the whole heart tube but expression becomes restricted to the atrial chamber from day 12. This atrial specific expression pattern turns the gene into a useful molecular marker for cardiac chamber specification. Furthermore the regulatory sequence of the gene will provide clues on atrial specific transcription regulation. Here we report the identification of the murine genomic sequence of the MLC-2a gene. In addition we compared genomic structure of the gene to other myosin light chains to show evolutionary structure conservation.

Methods: DNA-filter hybridization was used to identify genomic clones of the MLC-2a gene in a mouse genomic phage library. Clones were analyzed through restriction analysis and Southern hybridization and subcloned. To localize the transcription initiation site the rapid amplification of cDNA ends was used. Selected DNA regions were sequenced in both directions. Computerized sequence comparison was performed to search for homology to previously cloned genes. In addition sequence comparison of DNA and amino-acids was performed with previously reported MLC-2 sequences.

Results: Two phage clones were identified and purified. The phage with the largest insert size was analyzed in detail. The entire 5' flanking region was identified and sequenced. In addition, the exon-intron boundaries and 3' flanking region were determined. Sequence comparison revealed the presence of the final exon (11) of the mouse glucokinase gene 2.0 Kb upstream of the MLC-2a transcription start site. Genomic structure comparison of various myosin light chains showed a striking conservation at the intron-exon boundaries and an increasing intron size in more complex mammalian species. At the amino acid level there is 95% homology between the human and the mouse MLC-2a sequence.

Conclusions: The atrial regulatory myosin light chain is a member of a highly conserved family of sarcomeric proteins. In the mouse the MLC-2a gene is positioned 2.0 Kb downstream of the glucokinase gene, and contains 7 exons. The murine MLC-2a is highly conserved both at the amino-acid and cDNA level compared to the human MLC-2a.

Introduction

At the first glance there are striking similarities in expression pattern comparing the essential (MLC-1) with regulatory myosin light chains (MLC-2) in adult mammalian species. The ventricular isoforms of both MLC-1(v) and MLC-2(v) are expressed in slow twitch skeletal muscle and within the adult heart expression is confined to the ventricular chamber.[1,2] The atrial myosin light chain isoforms MLC-1a and MLC-2a are initially detected throughout the developing heart but both become restricted to the atrial compartment before birth.[3,4] However, a closer look at the expression patterns of MLC-1 and MLC-2, during development reveals important differences. For instance regarding the ventricular isoforms, the MLC-1v is expressed in the atrial chamber until shortly before birth, whereas the MLC-2v is ventricular specific at all times throughout development.[5] The main difference between the murine atrial isoforms is the early downregulation of the MLC-2a in the ventricular chamber compared to MLC-1a (day 12 versus day 17), the continued expression of MLC-1a in Purkinje fibers in the adult heart and its transient expression in developing skeletal muscle.[6] The temporo-spatial changes in the expression pattern of the MLC-2a gene make its mRNA and protein excellent molecular markers for cardiac development *in vitro* and *in vivo*.[4] Furthermore the regulatory sequence of the gene could potentially be developed into a molecular tool for atrial specific *in vivo* gene targeting experiments. Recently the near complete cDNA of the murine MLC-2a gene was cloned.[4] In order to initiate MLC-2a regulatory DNA studies the genomic MLC-2a clone and the missing part of the cDNA, including the 5' untranslated region, has to be identified. Therefore the MLC-2a cDNA was used as a probe to isolate the murine genomic clone. In this chapter we report the cloning of the MLC-2a gene. The sequence of the 5' flanking region, and part of the 3' flanking region are included, together with the intron sequences. The chromosomal position of the murine MLC-2a gene was established. In addition the genomic structure of the gene is compared to myosin light chain genes from other species to establish evolutionary changes and conservation.

Method

Isolation of genomic clones of the murine MLC-2a gene

The murine clone of the MLC-2a gene was isolated from the 129SVJ mouse genomic library in Lambda FIX®II Vector (Stratagene, La Jolla Ca). This library was constructed using DNA derived from hepatic tissue and has an average insert size varying form 9.0 to 22.0 Kb. The cloning procedure in brief; Y1090 R^- *E. coli* bacteria were grown overnight in the presence of 0.2% maltose and incubated

with a dilution of the library at 20.000 pfu per 600 µl cells for each 90 mm NZY plate. The infected bacteria were mixed with melted topagarose and immediately plated. Plates were incubated at 37°C until lysis was evident (± 8 hours). From each plate duplicate filters (Hybond-N, Amersham, Buckinghamshire UK) were lifted. The filters were treated by alkaline solutions to denature DNA, and were UV cross linked at a wave length of 312 nm. for 3 minutes. Filters were prehybridized for 30 minutes in SSPE, (0.9M NaCl, 0.05M Sodium Phosphate, 5mM EDTA), Denhardt's solution (0.1% bovine serum albumine, 0.1% Ficoll™, 0.1% Polyvinylpyrrolidone) and 0.5% SDS. Hybridization was performed using 400 bp of the murine MLC-2a cDNA as a probe.[7] For the hybridization 500.000 cpm/ml of heat denatured probe was added with 100 µg/ml of sonicated denatured salmon sperm DNA. Hybridization was performed overnight at 65°C and followed by high stringency washes going to 0.1x SSPE and 0.1% SDS at 65°C. The filters were wrapped in SaranWrap and positioned in a cassette with Kodak X-Omat (Rochester, NY) film for autoradiography. After three rounds of screening pure phage colonies could be selected and amplified. For amplification six plates were prepared for each clone, containing LB bottom agar enriched with 0.3% glucose, 0.075 mM $CaCl_2$, 0.004 mM$FeCl_3$ and 2mM $MgSO_4$. Approximately 10^5 pfu of phage was incubated with 0.1 ml Y1090 R^- cells for 20 minutes at 37°C, mixed with 7 ml melted topagar. Phages were harvested in SM buffer supplemented with 1/20 volume chloroform. Bacteria and phages were separated through low speed centrifugation. After a second high speed centrifugation step, supernatant was discarded and the glassy pellet containing purified phage was resolved in 1 ml SM buffer.

Characterization of genomic clones of the MLC-2a gene

All the primers listed in the methods section are shown in the 5'-3' orientation.

Candidate clones were characterized by restriction analysis. The restriction fragments were size fractionated by gel electrophoresis and transferred to a Hybond-N (Amersham) filter by capillary transfer. Filters were subsequently used for Southern analysis. To identify the 5' end of the cDNA an oligonucleotide was generated complimentary to GTTGTGCCTGCTTGGT (39-55), and for the 3' end the oligonucleotide GGCTCACCTCAATAAA (545-566)[4] Oligonucleotides were end-labeled by the enzyme T4 kinase in the presence of γ^{32}P-ATP. In order to identify the transcription initiation site of the MLC-2a gene, primer extension and the 5' rapid amplification cDNA ends (RACE, GibcoBRL, Gaithersburg MD) kit was used. The template for these reactions was poly A purified murine mRNA from adult mouse atrial tissue. RNA was isolated by the RNAzol method (Cinna Biotecx, New York NY) and poly A selected by using a poly-T column. Of the isolated mRNA 1 µg was used for the 5' RACE reaction to identify the transcription start site.

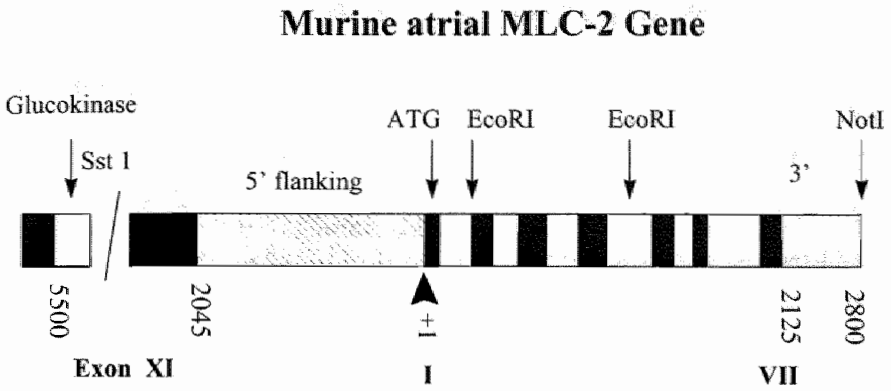


Figure 1. Schematic representation of the genomic structure of the regulatory atrial myosin light chain. Exons are shown as black, introns white and flanking regions as shaded boxes. Indicated is the position of exon 11 of the mouse glucokinase gene. The 5' flanking region is 2045 bp in length. Seven exons were identified, together with the introns the length of the genomic fragment is 2125 bp. The shaded box labeled 3' indicates the part of the 3' flanking region of which the DNA sequence is known.

Sequencing

Sequencing was carried out according to the method of Sanger et al.[8] Cycle sequencing was performed with ABI Prism Dye Terminator Cycle Sequence Ready Reaction kit (Perkin Elmer, Norwalk CT). The annealing temperature was adjusted to the calculated melting temperature of the oligonucleotide. The clones were partially sequenced by a dideoxynucleotide termination technique with α^{35}S-dATP using the enzyme Sequenase (GibcoBRL).

Results and discussion

Completion of cDNA

Two different phage colonies (AM1, AM2) were selected after 3 rounds of plaque purification. The phage genomic DNA insert size turned out to be 14 (AM1) and 13 Kb (AM2). For restriction analysis the purified phage DNA was incubated with BamH I, Bstx I, EcoR I, Hind III, Not I, Pst I and Sst I. Restriction analysis revealed a major overlap between the inserts of AM1 and AM2. Accordingly clone AM1 was used for further analysis. By Southern blotting the localization of the MLC-2a cDNA on the AM1 clone was confirmed. In addition by using oligonucleotides the localization of the 5' and 3' ends of the cDNA on the AM1 clone, were identified (not shown).[4] The genomic structure of the MLC-2a gene is shown in figure 1. Three restriction fragments, 6.6 (EcoR I-Sst I), 0.9 (EcoR I-EcoR I), and

1.5 (EcoR I-Not I) Kb of the phage genomic insert were subcloned into pBluescript II SK + (Stratagene, La Jolla, CA) for further analysis.
The sequence and position of the 5' end of the cDNA (figure 3) and 5' untranslated region were identified through primer extension and the 5' RACE method. An antisense oligonucleotide was made complimentary to MLC-2a cDNA (ATGAGAAGCTGCTTGAATC, 390-407).[4] Primer extension failed to present the complete cDNA end, as the extension product terminated in the first coding ATG triplet. Therefore 5'RACE was performed and the amplified DNA was cloned into TA-cloning vector (Invitrogen, San Diego CA) and sequenced. The first exon was identified, and the start of the coding region was designated as +1 (figure 1). The first exon appeared to be 22 basepairs (bp) long and terminated in ATG, providing the first triplet of the coding region (figure 2).

Genomic structure

Following determination of the transcription start site more then 3.0 Kb of the 5' flanking region was sequenced. Sequence comparison with known sequences in GenBank release 83.0 (1994) using the nblast NIH program, revealed a 96% homology of the region -2.5 to -2.0 Kb to cDNA (exon 11) of the rat glucokinase gene.[9] More recently the sequence of the murine glucokinase gene was reported and as expected homology to exon 11 was 100%.[10] The mouse glucokinase gene has been mapped to mouse chromosome 11, and therefore by inference the MLC-2a gene must be on the same chromosomal locus. Low homology (< 10%) to several genes, was found by the nblast computer algorithm for the 5' flanking region of the MLC-2a gene close to the transcription initiation site. These findings indicate that the 5' flanking region of the MLC-2a gene spans 2045 bp. In addition the sequence of the 3' flanking region and introns was determined (figure 2). AM1 contained approximately 700 bp of 3' flanking region. By sequence analysis of the genomic clone and comparison to the cDNA sequence the intron-exon boundaries were defined. The coding region of the MLC-2a gene encompasses 7 exons.

MLC-2 genomic sequence comparison

The Clustal software program version 1.6 (EMBL, Heidelberg Germany) was used to determine the homology between the human, murine and rat MLC-2 at the amino-acid level.[11-15] The amino-acid sequences are shown in figure 3. The calculated homology is 95% between human and mouse MLC-2a protein. The homology of the human MLC-2a to the other sarcomeric light chains is high (>50%) and drops to 47% when comparing the murine MLC-2a to the human smooth muscle isoform (MLC-2sm).

The genomic structure of the murine MLC-2a gene was compared to MLC-2 sequences from Drosophila melanogaster[16], Caenorhabditis elegans[17] and Homo sapiens ventricular MLC (Genbank file numbers M11947, M23366 and L01652 respectively). The genomic sequences have been published for a limited number of MLC isoforms. The evolutionary development of the myosin light chain genes is shown in figure 4. There is an increase in size of the gene, size of the introns and number of exons going form C. elegans to man. The number of exons is similar for the human MLC-2v, rat MLC-2v (not shown)[12] and mouse MLC-2a. Striking similarities are found when comparing exon-intron boundaries (figure 5). For instance the intron 2 with exon 3 boundary is basically identical in Drosophila (MLC-2), human (MLC-2v) mouse (MLC-2a\v) and rat (MLC-2v). Intron 1 is introduced after the first ATG in Drosophila, mouse and man, but not at this position in C. elegans. However the sequence of the C. elegans at the first ATG (ACATGGTACA) is a perfect site for the introduction of an intron. Whether the different splice sites require different spliceosome associated proteins is unclear. Sequence comparisons, as shown here, can be used to reconstruct the evolutionary tree.

With regard to the mammalian MLC-1 genes, the evolutionary origin of the human essential atrial and ventricular myosin light chains (MLC-1a and MLC-1v) is dated to ± 300 million years ago. This estimate is based on the fossil record and the divergence of chicken and human MLC-1 genes.[18] The fossil record indicates a separation between birds and mammals at that time point in evolution. The MLC-1 genes have a high homology at the DNA and protein level within the MLC-1 family.[4] Furthermore the genomic structure of the mammalian MLC-1 genes is characterized by the presence of 7 exons and positioning of the first intron within the first coding triplet (ATG).[1] The homology in expression, and the conservation of genomic structure and sequence suggest a separation of the cardiac MLC-1 and MLC-2 genes at a similar evolutionary timepoint.

The myosin light chain 1 and 2 genes belong to a family of genes encoding calcium binding proteins. More distant family members are the calmodulin and troponin-C gene. Calmodulin is a ubiquitous protein involved in the regulation of a wide variety of enzymes and in the transition of G1 to S phase of the cell cycle. The calmodulin mRNA (chicken) is composed of 8 exons and the second exon terminates in ATG, the first coding triplet, similar to the MLC-2 genes (figure 1).[19] This structural homology points in the direction of a mutual ancestor.

-2070 TTAAATATTT AAAACTACTT AAACAGGCCT GCTGATTGTT CTTTGGCCTC TGTGCATACT GATGGTGGTG
-2000 AAGACATACG AGAAACGTCA CTGATATACN CAGACACTCC CTAAACTCGG CCATGTCTCA ATTTGCTGTG
-1930 TCCTCCAGAC ACTATACCCA GGGGCCTGTG AGAGGCTAAC TATAGGAATA GGAATGTGGA GGCATGGGGA
-1860 CACTAGGGGT ACTGAATAGC CCCTGCTGGC CCAGACCCCA CCCCCACCCC CAGCTGCCCT TGCAACATGT
-1790 GAAAGATAGA GGCTAGGATG CCAGGCTAGC TGCCACTCTG AGGAGGGGAG GGGTTGGTCC AGCTCTAGGA
-1720 TGGTCTCAGC AACATTCTAT TCTCAGAACC AGGGAGGAAG CAATATGGGC CAAGGCTAGG AGCTATTAAA
-1650 CAGTTTCTAA GGTCCCTCCA GAGGCTCTGG TAGATGGAAT CTACCTAATN CTTTTCTATG CAGCATGAAT
-1580 ATGCCTTCAG GTGAACGGAC TTGTATGCGT TTAATGCACA TTTAATGCAC ATTTATCTCT CACTGGGAAC
-1510 CCAGGGAGTC AAAGGCTTAC TCTGTTCTCT GCTTCAGAAA AAAAATCCCA AATTGGTCGG TTCTGTGGGA
-1440 AAGTCACATT TTTCTTGAAA TCTTGCAGTT TTCTGGTAAT TCTCGTGTAG AATTGCTTCT ATCGGCTCAC
-1370 TTTCTTGAAC CAACACGAGC ATTTGCATTA ACAGAGTACA ACACTCATAT TGGCACACTC ACACATTGGA
-1300 ACAGATTTTC AGTCTTTCAT CTTTTCGTAC ATACTGGGAT CTGAGGCATT GTGCTGGTTT GTTTGTTTGT
-1230 TTTGTTTTGC TCTCTTTATA CATCCAACTT GAATGATTGA ACTTTGTTTT CAGGTAGTTA GGAGGGTGTT
-1160 ATAAAATATT CATACAAAAG GAGGCCTTGG AGTTGTGTTG TGGTACAGAC TTGTAATCCC AGCACCAGCC
-1090 TGGAACATCG TGAGTTCCAG GCTAGCCTTN GCTACATAGC CGACCCTATC TCAAAATAAA AAGAGGATGG
-1020 AGTGGACTTG AGTCCTACTA ACAGGAGATA GTGCCCATAT CCTAACTCAC CTTGCATAGC CACATTGGGG
- 950 TGAGGATGAA CACTGAACCT AGTCAGAGAC CTACTCATGA TCTCCCCTAA GCTGTCCAGC TTCTCTGACT
- 880 CCTCTGCCAC ACTAAGGTCC TCAAAAGTTG ACGGCTGCTG CTTGCAACCA AAGCCCCTTG GTCCCAAATT
- 810 CAGAAACATG AGATGGTCTT TGTGGGCCAG TTTCATAAGA ACCCAAAACT TTTCCTGTGA AACAGCAATC
- 740 ATGTAGGCCT GCCACGTGCC AGGGTTTCTG ATTCCTGATG TGGAGCAAGC TTGGGTGCCC AGGAGAAGGC
- 670 TACAGCTCTC AGTTGGGAAC GTCAGAGAGA ATTTGGAGGA GGAGGGTAGA GTTGAGATGG AGCTTAAGAA
- 600 ACACAAGATT CCGTATAGGG TGGGTACGCG CGTCTCAGCA GGACAGGCCG AGTGAAGAGC ATGTCTGAAT
- 530 ATGTAGGGAG AGCAGCATCG GTCACTCCAA TCCTGAGTTA CCCCCAAATC AAGCCAATGG TCCTGTGGGG
- 460 CTCACACGCC TGGTTAGAAC TGCTGTCTGT CCATGTCAGA TGGCCAGCAG GCCAGAACTC AGGGGTTTTA
- 390 CTTGACCCAG GGAATTCAGG CCAGAGACCA TTCTCAGAAT CATGTACTGC TTCCCAAGCC CAGAGTATAG
- 320 ACAGCACTAG TCAGGTGAGA GGACGGTCAG ATGCAGGCAT CACAGGGAAT CACACCACCA TGACCTGGAC
- 250 TAGCACACGT TTGGAGCGTC TCTGACCAGG GAAGACATGG AGCCTGGTCA CAAGAGATGG ATGGGTGGAT
- 180 CCTAGGATGT TACACTTGGG GGCCTGAAGG TTACCAGGAG TGTGAATAGG ATCTATGTGG AGCCCTGTCT
- 110 TTCTCCACCA GCGGGGGCTT CCAAGTTTAG CCCAATCATC CATTCCTCTC CTTATTTGGG CCTCTGGGTG
- 40 ATAAGGCTGA GACATAAAGA GCTGGGTGGC AGGATCACTC ATGCAGACCG CCACAGATCA TGGTGAGTGC
EXON
+1
I
22
+ 31 TGGGTGGGGA ACAAGAGACC AGCCACACCC AGGTGCTAGT CTCTGAGCTG CCCTGGGACA GAGCCAGGGA
+ 101 GCTTGTGTGC CTGTGCTGGG CTCAGGGAAG TGATCTCGGG TGGGGAAAGG CACCTCTCCC TGGCACTCCC
+ 171 ACGAGACTAC ATGAGGTAGG GAGGAGGTCT TTGGAGAACT GGAGAGAGGA CCCCTCTGAT CATCAGTATT
+ 241 ACAGGCCAGT AGGAAGGCTG GGACCCGGCA AGCTGCAGCC ACCAAGCAGG CACAACGTGG CTCTTCTAAT
EXON 23
****
****
+ 311 GTCTTCTCAA TGTTCGAACA AGCCCAGATT CAGAATTCAA GGAAGTGAGT AGCCTATCCC AAACAGCCTC
II
123
+ 381 TCCCCGATCC GGGGAGCCAC TTCATACCCC ATGCACACGG CTCTCCTCTG AGGGTCAGTG CTCCTGCCTC
+ 451 CAACCTCGCT CATACTACCC GAGACTCTCC CTTACCCCGC CATCCCTGGA AGAGATGGGG TGAGCAGCTC

Figure 2.

```
+ 521 CCACCTGCAG GCCATCTCCT CACAGGCCTT CAGCTGCATT GACCAGAACA GGGATGGGAT CATCTGCAAA
 EXON                          124                                    III
+ 591 TCAGACCTGA AGGAGACCTA TTCCCAGCTC GGTGAGTGGT TCCAGGAGCA ACCTCCCCAG GACCCGGCTT
                                      200
+ 661 TGTTTACCTT CAGGCACTTG TTCCGAGAGA GGACAGTGCC CCGCCCTTTT CTAACTTAGC TTCATTTCAG
+ 731 GCAGGAATTC CCCTCCCCCA CCTGCCACAC AGCTAGAGCT GCGCCCGTGC CTCTGCCCTC ACTGACNTGG
+ 801 AAGTCCCGGG AGGGCGGCCT GGGTCACCTC AAGACCTACT TTCCTAGGGA GGGTAAGTGT TCCGGAGGAA
 EXON                                                  201
+ 871 GAGCTGGACG CCATGCTGCA GGAAGGGAAG GGTCCCATCA ACTTCACCGT CTTCCTCACA CTCTTCGGGG
                                    IV
+ 941 AGAAGCTCAA TGGGGAGTTG GCAGGGAGGT NTGAGTCTCT TAGCCAGCGC AACCGACTCA CAGACCCGTT
                306
+1011 NCCCACGTGA GGTCACCTAA CTCCCCTGAA ACCTGTGCAC CCTGCCTGCC CAGGGCCAAN GCAGAGGATG
+1081 TAAAGCGCCT GAACTGGCTT GGGTTCAGGA AATATCCACT TCCTAATCAC AGTTTTGACT TCCCAGCCGG
+1151 CCTGCCAACC CCTTTGATGC CCAAGAAACA CTCTGAATTC CGAATTCAAG TTTCTTTGCA CCCTCCAATT
+1221 TAACCCCACC CCACCTTTAT TTCAATTCGG TATAGCCGGA GCCCTTTCAG GCCCCAGTAA CAAAGAAGGC
+1291 TTCCCTTGTA ACTTCCAAGT TTCATTCCTG GGTGGGGCCT TCGCGGCTGG GGATTCTGAG AGTTTCAACT
+1361 TGGCTGGCCT TGGCCCGCCC CCCCTTCCAG GGACGGACCC CGAGGAAGCC ATCCTGAGTG CCTTCCGGCA
 EXON                          307
+1431 TGTTTCGAAC CCCAGCGGCC AGGGGGTGGT GAACAAGGAA GAGTAAGTGA GGCCCTTTCA TGCCCCAGGG
          V                                  390
+1501 GCGGGGCCAT GGGGAAACCC AGAAGGTGGT GAGGTTCTTT GTGCGCCTCA GGTGGGTGAG CAGCTCTACC
+1571 CTGGGCCCCA GCCTCGCACT CCCAGCTCAG GCATTCTCTC TCCTCAGGTT CAAGCAGCTT CTCATGACCC
                                    391*** ********** *****
+1641 AGGCAGACAA GTTCTCTCCT GCTGAGGTGA GACACCCAGA CCCACCAACT CCTGTCATGA CCGTGCCCCC
 EXON      VI                  440
+1711 CCTCACGCCC CACAGGTGTG TTGGTNGGAG CCNAATGGAT AACAGGATCT TCTGAGAAGA GATATAGCTC
+1781 TGTAACCCAA AGGAGACGCA GATCCAAGAG GTGGTTATTT ATGTGGTGAG CCTTGGCCTT CCCCATTGGC
+1851 TGTTCTATNT GGTTCAGAAC AGGAAATCCA GCANCTGCAG TTTGGGGTTG AAAAGGCTGA GTCCGGAAGC
+1921 ACAGCTGTTG GGNACAGGGT ATCATAGGGT GACATCTCTG CTNGTCCCCT GCAGGTAGAG CAACTGTTTG
 EXON                                                  441
+1991 CGCTGACACC CATGGACCTG GCCGGCAACA TTGACTACAA GTCTCTCTGC TACATCATCA CCCACGGGGA
                                  VII                  ▼
+2061 TGAGAAAGAG GAGTAGACAG GCTGAGGTCA CCTCAATAAA CTCTGTGCCT GAACTAGAAC GAGGTGTTGN
                          * ********** **********                    587
+2131 GCTTGACTGT GGCAGATGGG CTGGAGAGGG ACCCCACATG GTACCACCTT GAGTCACCCT ACACCACCAA
+2201 GTAAAATATG AACCTCTATT TGGCCCACGG GTGAACTCCA AGAGAGGAAA CGGGGTACAC ACAACCTCCA
+2271 TGAGCCCAGA ATGAGCCACA AGTGGGCAAC CCAGCTCTTG ATGTCTGGGG AAGGAAGAAA CATCTTGCTG
+2341 GGACCCTTTC TGCCCAACAT CCAGCTAGAG GATGCAGCCC CAGGGACTCC CAACCACATA GCCTTCCCGT
+2411 GCACTCAGGT CTACTGCCCA CTCTTCTCTT GATGTGCTGC CTGTGTGGGC CTATGGGAAA ACCACAGGCA
+2481 GAACTCAATG CTGGCTCTGC GGTGCAAGCC TTGCCTAGCA TGTTCTATGT CCTGGATTCC ATTCCCAGCA
+2551 AGCACCTCTC CCTTAGGTGG GGGAATCACA ATCATATTGG GATTATAGGA AGATTTTGAC ATGATGGGAC
+2621 ATGATGCTGG ACGCAGAGCC AGATCGAGTC GACGCCCTAT AGTGAGTCGT ATTAGAGCTC G
```

Continued figure 2. The sequence of the genomic clone of the regulatory atrial myosin light chain. Indicated are poly adenylation signal on the final exon of the glucokinase gene (arrow). The exons of the MLC-2a gene are underlined and the cDNA is numbered (italic numbers). The oligonucleotides used are indicated by (*). The initiation and termination sequence of the MLC-2a gene are indicated by arrows.

Myosin light chain isoform	Amino acid homology (%) compared to Mouse MLC-2a
Human MLC-2a	94.9
Human MLC-2v	59.4
Mouse MLC-2v	62.8
Mouse MLC-2f	57.1
Rat MLC-2f	58.2
Human MLC-2sm	47.4

Figure 3A. Mathematical determination of amino-acid homology. The homology of the various MLC-2 genes compared to the murine MLC-2a gene is shown.

Sequence comparison of the human regulatory myosin light chain with other known human regulatory light chains and rat MLC2f.

```
MASRKAGTRGKAAATKQAQRGSSNVFSMFEQAQIQEFKEAFSCIDQNRDG   MouseMLC2a
..........V.....................................     HumanMLC2a
..PK..KK.AGG.--------N.........T.........TIM......   HumanMLC2v
..PK..KKRIE-------GGT...-.....-T.........TIM......   MouseMLC2v
..PK..KR.AGAE-------.........D.T.........TV.......   MouseMLC2f
..PK..KR.--A..E-----.......D...T.........TV.......   RatMLC2f
.S.KR.KA--.-.TK.RP..AT....A..D.S.........NM.......   HumanMLC2sm

IICKSDLKETYSQLGRVSVPEEELDAMLQEGKGPINFTVFLTLFGEKLNG   MouseMLC2a
....A..R.......K................................     HumanMLC2a
F.D.N...D.FAA.-R.N.KN..I.E.IK.AP.........M.....K.    HumanMLC2v
F.D.N...D.FAA..R.N.KN..I.E.IK.AP.........M.....K.    MouseMLC2v
..D.E...D.FAAM.RLN.KN......MK.AS.........M.....K.    MouseMLC2f
..D.E...D.FAAM.RLN.KN......MK.AS.........M.....K.    RatMLC2f
F.D.E..HDMLAS...-NPTD.Y.EG.MS.AP..Y...M...M.......   HumanMLC2sm

TDPEEAILSAFRMFDPSGQGVVNKEEFKQLLMTQADKFSPAEVEQMFALT   MouseMLC2a
.................K.....D......L.............L.T..   HumanMLC2a
A....T..N..KV...E....LKA.YVREM.T...ER..KE..D....AF   HumanMLC2v
A....T..N..KV...E...SLKA.YVREM.T...ER..KE.ID....AF   MouseMLC2v
A...DV.TG..KVL..E...TIK.Q.LEE..T..C.R..QE.IKN.W.AF   MouseMLC2f
A...DV.TG..KVL..E...TIK.QFLEE..T..C.R..QE.IKN.W.AF   RatMLC2f
....DV.RN..AC..EESS.FIHE.HLRK..T.MG.R.TDE..DE.YREA   HumanMLC2sm

PMDLAGNIDYKSLCYIITHGDEKEE                             MouseMLC2a
.........................                             HumanMLC2a
.P.VT..L...N.VH.....E..D                              HumanMLC2v
.P.VT..L...N.VH.....E..D                              MousMLC2v
.P.VG..V...NI..V.....A.DQE                            MouseMLC2f
.P.VG..V...NI..V.....A.DQ.                            RatMLC2f
.V.KK..FN.VEFTR.LK..AKDKHD                            HumanMLC2sm
```

Figure 3B. Comparison of the amino-acid sequence of the human MLC-2a, MLC-2v and MLC-2sm, with murine MLC-2a, MLC-2v, MLC-2f and rat MLC-2f. Amino-acids are denoted using the single letter code. Identical amino-acids are indicated by dots (.), dashes (-) indicate spacing to optimize homology.

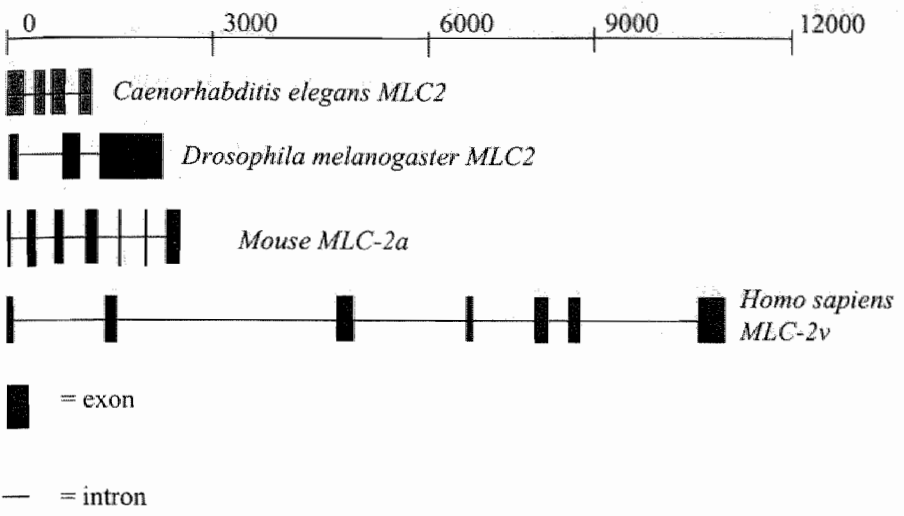


Figure 4. Schematic representation of genomic structure of MLC-2 genes from various species. Indicated as boxes are the exons, the introns are represented by a line. The size of the introns and exons can be derived form the scale (in basepairs) at the top. Sequences are derived from Genbank: M23366 *C. elegans*, M11947, *D. melanogaster* L01652: H. sapiens

Intron-Exon boundaries of regulatory Myosin light chain genes in different species

	C.elegans	D. melanogaster	R. norvegicus MLC-2v	H. sapiens MLC-2v	M.musculus MLC-2a
e1 - i1	acaag gttag	ccatg gtaag	ccatg gtgag	ccatg gtagg	tcatg gtgag
e2 - i2	gacag gtagg	cgaag gtgag	aggag gtgag	aggag gtggc	aggaa gtgag
e3 - i3	acaag gtaaa		cctcg gtatg	ccttc gtatg	gctcg gtgag
e4 - i4	cctaa acatc		tacag gagct	taagg gtaag	caatg ggagt
e5 - i5			tacag tgtcc	gatta gtaag	aagag taagt
e6 - i6		aagaa cgctc	tgcag atcga	aggag gtaat	ctgag gtgag
i1 - e2	ttcag gaact	gaccg agtag	tccag tcacc	cccag gcacc	tacag gccag
i2 - e3	ttcag gaact	gatag gcctt	tacag gcctt	cacag gcctt	cacag gcctt
i3 - e4	ttcag ggaaa		gccag gacga	ggcag gagtg	cctag ggagg
i4 - e5			tacag gagct	ttcag gagcg	tccag ggacg
i5 - e6			tacag tgtcc	tgcag cgttc	ctcag gttca
i6 - e7			tgcag atcga	tgcag gttga	tgcag gtaga

Figure 5. Intron-exon boundary comparison of MLC-2 sequences from various species. Ten bases surrounding each intron and exon boundary are shown. Listed are the MLC-2 genes from *C. elegans* and *D. melanogaster*, MLC-2v: *R. norvegicus*, MLC-2a: *M. musculus*, MLC-2v: *H. sapiens*. In bold the conserved bases in rat, man and mouse MLC-genes are indicated. Note the introduction of the first intron behind ATG (not in C. elegans) and the striking similarity in the intron 2- exon 3 boundary (not in C. elegans).

Conclusions

The atrial regulatory myosin light chain is a member of a highly conserved family of sarcomeric proteins. In mouse the gene is positioned 2.0 Kb downstream of the glucokinase gene on chromosome 11, and contains 7 exons. The murine MLC-2a is highly conserved both at the amino acid and DNA level compared to the human MLC-2a. The identification of the transcription start site and the cloning of the 5' flanking region of the MLC-2a gene, provides the tools to study transcription regulation of the MLC-2a gene in physiological and pathological conditions. Furthermore the regulatory DNA of the gene, once characterized, can be used in atrial specific gene targeting experiments in mice.

Literature

1. Barton PJ, Cohen A, Robert B, Fiszman MY, Bonhomme F, Guenet JL, Leader DP, Buckingham ME: The myosin alkali light chains of mouse ventricular and slow skeletal muscle are indistinguishable and are encoded by the same gene. *J Biol Chem* 1985;260:8578-8584
2. O'Brien TX, Lee KJ, Chien KR: Positional specification of ventricular myosin light chain 2 expression in the primitive murine heart tube. *Proc Natl Acad Sci U S A* 1993;90:5157-5161
3. Price KM, Littler WA, Cummins P: Human atrial and ventricular myosin light chain subunitsin the adult and during development. *Biochem J* 1980;191:571-580
4. Kubalak SW, Miller-Hance W, O'Brien T, Dyson E, Chien K: Chamber-specification of atrial myosin light chain-2 expression precedes septation during mouse cardiogenesis. *J Biol Chem* 1994;269:16961-16970
5. Schiaffino S, Samuel JL, Sassoon D, Lompre AM, Garner I, Marotte F, Buckingham M, Rappaport L: Non-synchronous accumulation of a-skeltal actin and b-myosin heavy chain mRNA during early stages of pressure overload-induced cardiac hypertrophy demostrated by in situ hybridization. *Circ Res* 1989;64:937-948
6. Whalen RG, Butler-Browne GS, Gros F: Identification of a novel form of myosin light chain present in embryonic muscle tissue and cultured muscle cells. *J Mol Biol* 1978;126:415-431
7. Poetter K, Jiang H, Hassanzadeh S, Master SR, Chang A, Dalakas MC, Rayment I, Sellers JR, Fananapazir L, Epstein ND: Mutations in either the essential or regulatory light chains of myosin are associated with a rare myopathy in human heart and skeletal muscle. *Nat Genet* 1996;13:63-69
8. Sanger F, Coulson AR: A rapid method for determining sequences in DNA by primed synthesis and DNA polymerase. *J Mol Biol* 1975;94:444-448

9. Andreone TL, Printz RL, Pilkis SJ, Magnuson MA, Granner DK: The amino acid sequence of rat liver glucokinase deduced from cloned cDNA. *J Biol Chem* 1989;264:363-369
10. Postic C, Niswender KD, Decaux JF, Parsa R, Shelton KD, Gouhot B, Pettepher CC, Granner DK, Girard J, Magnuson MA: Cloning and characterization of the mouse glucokinase gene locus and identification of distal liver-specific DNase I hypersensitive sites. *Genomics* 1995;29:740-750
11. Hailstones D, Barton P, Chan-Thomas P, Sasse S, Sutherland C, Hardeman E, Gunning P: Differential regulation of the atrial isoforms of the myosin light chains during striated muscle development. *J Biol Chem* 1992;267:23295-23300
12. Henderson S, Spencer M, Sen A, Kumar C, Siddiqui M, Chien K: Structure, organization, and expression of the rat cardiac myosin light chain-2 gene. Identification of a 250-base pair fragment which confers cardiac-specific expression. *J Biol Chem* 1989;264:18142-18148
13. Nudel U, Calvo JM, Shani M, Levy Z: The nucleotide sequence of a rat myosin light chain 2 gene. *Nucleic Acids Res* 1984;12:7175-7186
14. Kumar CC, Mohan SR, Zavodny PJ, Narula SK, Leibowitz PJ: Characterization and differential expression of human vascular smooth muscle myosin light chain 2 isoform in nonmuscle cells. *Biochemistry* 1989;28:4027-4035
15. Dalla Libera L, Hoffmann E, Floroff M, Jackowski G: Isolation and nucleotide sequence of the cDNA encoding human ventricular myosin light chain 2. *Nucleic Acids Res* 1989;17:2360
16. Parker VP, Falkenthal S, Davidson N: Characterization of the myosin light chain 2 gene in Drosophila melanogaster. *Mol Cell Biol* 1985;5:3058-3068
17. Cummins C, Anderson P: Regulatory myosin light chain genes in Caenorhabditis elegans. *Mol Cell Biol* 1988;8:5339-5349
18. Kurabayashi M, Komuro I, Tsuchimochi H, Takaku F, Yazaki Y: Molecular cloning and characterization of human atrial and ventricular myosin alkali light chain cDNA clones. *J Biol Chem* 1988;263:13930-13936
19. Simmen RC, Tanaka T, Ts'ui KF, Putkey JA, Scott M, Lai E, Means A: The structural organization of the chicken calmodulin gene. *J Biol Chem* 1985;260:907-912

Chapter

The in vitro characterization of the murine atrial myosin light chain 2 promoter

Pieter A. Doevendans, Heleen Erkens, Ronald Bronsaer, Pilar Ruiz-Lozano, Kenneth R. Chien, Marc van Bilsen.

Abstract

Background: The mRNA of the atrial myosin light chain 2 was recently reported to be expressed exclusively in atrial tissue. Initially expression was seen in both cardiac chambers but expression becomes restricted to atrial cardiomyocytes early in embryonic mouse development (day 12/21). Accordingly, the regulatory DNA of this gene must contain DNA elements required for atrial specific gene expression. In this study it was attempted to identify the DNA elements that mediate atrial myocyte specificity.
Methods: Deletion fragments of the atrial myosin light chain promoter were cloned into luciferase reporter gene containing vectors. Different cell types were transiently transfected with the promoter constructs. After 48 hours cells were lysed and the luciferase activity was measured. The importance of a DNA element (Initiator) involved in transcription initiation was studied by mutation analysis.
Results: The 0.2 Kb core promoter fragment was active in cardiomyocytes and AT1 cells, whereas low level of activity was seen in COS cells and skeletal muscle cells. A 0.45 Kb promoter fragment was even more active in cardiac myocytes, but was silent in other cell types. Extension of the MLC-2a promoter (>0.9 Kb) lead to reduced luciferase levels in cardiac myocytes, indicating the presence of a silencing sequence upstream of 0.45 Kb. Promoter activity was markedly reduced by complete or partial deletion of the conserved initiator site.
Conclusion: The regulatory DNA of the atrial myosin light chain 2 gene mediates cardiac specific expression. The short promoter fragments (0.45 Kb) revealed high levels of reporter gene expression compared to the longer fragments. The Initiator site of the myosin light chain 2 gene was shown to be crucial for transcription. No atrial specific regulatory elements were identified with transient transfections in cultured rat neonatal atrial and ventricular cardiomyocytes.

Introduction

The cDNA of the murine atrial myosin light chain 2 (MLC-2a) gene has recently been cloned.[1] The MLC-2a belongs to the family of calcium binding proteins and can be phosphorylated. In striated muscle the protein modulates contraction upon activation. The protein is part of the hexameric myosin unit. Each myosin unit has two paired light chain isoforms. The MLC-2a or regulatory light chain is located at the site of the hinge of the myosin heavy chain molecule. This hinge is responsible for the movement within the heavy chain molecule during contraction.

The cDNA was used to study the expression pattern of the MLC-2a gene during embryonic development and in the neonatal and adult stage. The atrial MLC-2 gene is expressed from day 8 of embryonic development in the whole heart tube, but becomes restricted to the atrial compartment from day 12. The initial expression seen in the ventricular chamber is down regulated sharply at day 12 post coitum. In the growing ventricle the atrial MLC-2 is replaced by the ventricular isoform (MLC-2v).[2] The MLC-2v isoform delineates ventricular chamber formation at all stages. The value of MLC-2a and v as molecular markers in chamber specification was shown *in vivo* and *in vitro*.[1-4]

To study the transcriptional mechanism, underlying atrial myocyte specific gene expression, we have identified the genomic clone of the murine MLC-2a gene and cloned the full length of the MLC-2a gene and the 5' flanking region, containing the essential promoter. In search for additional regulatory sequences the first intron and 3' flanking region of the gene were sequenced (chapter 6). Transcription activity in transient transfection assays was compared in rat neonatal atrial and ventricular myocytes, a stable atrial myocyte line (AT1), skeletal myoblasts and myotubes (C2C12) and monkey kidney cells (COS-7). The transcription initiation site turned out to be a conserved *cis* acting element. The role of this element in transcription regulation was studied by deletion and replacement mutation analysis. In addition the potential role of intron 1 was studied in transfection assays.

Here we report on the *in vitro* characterization of the regulatory DNA of the MLC-2a gene. The short promoter fragments are active in cardiomyocytes and AT1 cells, while low level of activity was seen in COS cells and skeletal myoblasts. An upstream silencer sequence results in strong repression for the longer promoter constructs. Furthermore we found markedly reduced activity by complete or partial deletion of the conserved Initiator site.

Method

Cell culture and transfection

Monkey kidney derived COS-7 cells (ATCC CRL-1651) were grown in Dulbecco's modified Eagle's medium (DMEM, Cat no 42430-025, GibcoBRL, Paisley UK) supplemented with 10% fetal bovine serum and 100 IU/ml penicillin and streptomycin at 100 ug/ml (Gibco).

Mouse C2C12 myoblast (ATCC CRL-1772), obtained from the American Type Tissue Culture Collection, (Rockville MD), were maintained in growth medium (85% DMEM, 15% fetal calf serum, 4 mM glutamine and 50 µg/ml gentamycin (Boehringher Mannheim, Mannheim Germany) in a humidified incubator at 5% CO_2. Cells were split at 50% confluence at a split ratio of 1:4 using 0.25% trypsin (w/v), 1 mM EDTA in $Ca^{2+}$and Mg^{2+}-free PBS. Myoblast transfections were performed at 50% confluency, in growth medium at day 1. For myotube transfections at 90% confluency medium was replaced by differentiation medium (DMEM plus 2% horse serum [Boehringer], 4 mM glutamine, 50 µg/ml gentamycin and 0.3 µM insulin [Boehringer]). For myoblast transfection regular growth medium was maintained.[5]

AT1 and AT2 cells were kindly provided by Dr W. Claycomb (New Orleans USA). The AT1 cells were injected subcutaneously in syngeneic recipient mice. After 10-12 weeks the subcutaneously growing tumor tissue was removed aseptically and minced into small fragments (< 2 mm^3) and incubated overnight in 0.125% trypsin solution at 4°C. The tissue was dissociated the next day by a series of 15 minutes treatments with a 0.1% solution of collagenase in serum free PC-1 medium (Ventrex, Portland ME) on a shaker at 37°C. The cells from the consecutive treatments were pooled and plated at a density of $12.5x10^4$ cells/mm^2 in 60 mm petri dishes coated with gelatin and fibronectin. The medium (PC-1) was supplemented with 10% fetal calf serum, 15 µg/ml of insulin, 1 nM dexamethasone, 100 IU/ml of penicillin and 100 µg/ml of streptomycin and was changed every other day. AT-2 cells were maintained in DMEM supplemented with 10% fetal bovine serum and 100 IU/ml penicillin and streptomycin at 100 mg/ml(Gibco-BRL).

To obtain primary cultures of cardiomyocytes enriched in atrial and ventricular myocytes respectively, the hearts from day 1-4 old neonatal rats were dissected and the atria and ventricles were separated by pulling off the atrial appendices. In separate rotation flasks, atrial and ventricular tissue was digested by repeated collagenase treatment and separated from fibroblasts by applying a discontinuous Percoll gradient.[6]

Transient transfection experiments were performed in COS, C2C12, AT1 and AT2 cells and in rat neonatal atrial and ventricular cardiomyocytes.[7] Transfection experiments were performed using the $CaPO_4$ precipitation method[8], except in

case of AT1 and AT2 cell transfections for which liposome mediated transfection was applied.[9,10]

For $CaPO_4$ transfections cells were cultured in 60 mm plates and 8 µg promoter-luciferase plasmid was used, with 2 µg of pON-249 plasmid DNA (ratio 4:1). pON-249 harbors the β-galactosidase reporter gene driven by the CMV promoter. RSV-luciferase was used as a positive control, as both viral promoters (RSV and CMV) are expected to be active in all cell types studied. For $CaPO_4$ transfections plasmids were mixed with 0.25 M $CaCl_2$, to which an equal volume of 2x BBS (50 mM N,N-bis (2-hydroxyetyl)-2-aminoethanesulfonic acid 280 mMol NaCl, 1.5mM Na_2HPO_4, pH 6.95) was added under continuous aeration. Concoction was left for 20 minutes at room temperature and added to the plates dropwise. Plates were incubated for 16 hours at 37°C, washed twice with medium and maintained for 48 hours with DMEM supplemented with glutamine, antibiotics and 10% serum. Cells were treated with lysis buffer containing (100 mM K_2HPO_4:KH_2PO_4 (92:8 vol), pH 7.9, 0.5% Triton x-100, 1mM dithiothreitol). In the lysate luciferase activity was measured on a luminometer (Perstorp Analytical Lumac, Landgraaf the Netherlands), whereas galactosidase activity was assessed by spectrophotometry (Pharmacia Biotech AB, Uppsala Sweden). Each transfection was performed using three independent cell cultures and by using three different cesium chloride purified plasmid preparations.

Approximately $12.5x10^4$ AT1 cells/mm^2 were plated in 60 mm dishes and transfected with 26 µl LipofectAMINE (GibcoBRL) containing 5.2 µg plasmid DNA keeping the promoter luciferase/pON-249 ratio constant at 4:1. No antibiotics were administered during transfection (6 hours). After 48 hours cells were lysed and luciferase and galactosidase activity were determined. [11-13]

Plasmid constructs

To create a convenient Hind III restriction site at the 3' end of the promoter a 428 bp PCR fragment was generated using sense (5' AGGCCAGAACTCAGGGGTTTTATC 3', position: -409,385) and antisense (5' AGCTTGATCTGTGGCGGTCTGACTGAG 3', position -3,+18) primers (for complete sequence see chapter 6). The PCR product was cloned into TA-cloning vector (Invitrogen, San Diego CA) named HEX1, and digested using an endogenous BamH I site and the added Hind III site, the fragment was isolated by gel electrophoresis and purified with Qiagen (Qiagen, Hilden Germany). The purified fragment was used to replace 200 bp at the 3' end of the promoter using the endogenous BamH I at position -203 of the MLC-2a promoter in pBluescript II SK + (Stratagene La Jolla CA). The sequence of the PCR product and the BamH I ligation were checked by two directional sequencing.

To determine the functional importance of the various regulatory elements, deletion fragments were generated using convenient restriction sites resulting

in fragments of decreasing length (Pst I: 3.0 Kb, Pvu II: 1.8 Kb, Ssp I: 1.2 Kb , Hinc II: 0.9 Kb, BamH I: 0.2 Kb). In addition the 0.45 Kb PCR fragment from clone HEX1 was used. The various regulatory DNA fragments were cloned into pXP1, which is a promoterless vector containing a multiple cloning site upstream of the luciferase reporter gene.[11,14]

To evaluate the role of the first intron in transcription regulation, intron 1 was amplified by PCR using sense primer (5' AAGCTTAGATCTGGGTGGGGAA-CAAGAG 3', position: +29,47) and antisense primer (5' GGATCCGGTACCTG-TAATACTGATGAT 3', position: +229,245). This PCR reaction created convenient restriction sites 5' (Hind III, Bgl II) and 3' (BamH I, Kpn I). The intron was cloned upstream of the 0.2 Kb promoter fragment in both orientations (5'-3' and 3'-5'). The orientation of the intron was checked by restriction analysis and by sequencing in both directions. As mentioned sequence analysis revealed a consensus transcription initiation site (Inr). To further characterize the transcription initiation site positioned from -3 to +5, four additional PCR fragments were generated using different antisense primers resulting in mutations (M1, M2), and partial (D1) or complete deletion (D0) of the transcription initiation site. The PCR products were cloned into TA-cloning vector and the BamH I-Hind III fragment was cloned into pXP1.

Sequence comparison and putative cis element identification

To identify putative *cis* elements the computer program Sigscan (MBCC, St. Paul MN) was used. The software is based on the biochemical characterization of *cis* elements and consensus sequences reported in literature. By comparing the *cis* elements identified by Sigscan to previously characterized elements contributing to muscle specific transcription regulation, we selected the potentially interesting sites.

Immunocytochemistry

For immunocytochemical analysis atrial cardiomyocytes were cultured on poly-D-lysine (Serva, Heidelberg Germany) precoated cover glasses. Cells were washed 3 times with PBS and fixed with methanol for Giemsa-staining or with 4% formaldehyde and 0.2% glutaraldehyde (all three from Sigma, St. Louis MO) in PBS for immunocytochemistry during 18 h at 4°C. For immunofluorescence labeling, fixed cells were washed 3 times with PBS and incubated for 30 min with affinity-purified polyclonal rabbit anti-mouse MLC-2a antibodies (1:500 v/v dilution) and monoclonal mouse anti-myomesin antibodies (1:400 v/v dilution, generously provided by J.C Perriard, Zürich), in PBS containing 5% bovine serum albumin, 0.1% Tween 20 (Sigma) and 4',6-diamidine-2'-phenylindole dihydrochloride (DAPI, Boehringer Mannheim, Mannheim Germany) at 37°C in a

humidified atmosphere. The antibodies were aspirated, and the cells were washed 3 times with PBS before incubation with Texas Red-labeled anti-mouse immunoglobin antibodies from goat (Amersham, Buckinghamshire UK) or fluorescein isothiocyanate-labeled anti-rabbit immunoglobin antibodies from donkey (Sigma), respectively. Slides were mounted with Gelvatol and 2.5% triethylenediamine.

Northern blotting

Total RNA was isolated from AT1, AT2, H9c2, HEK293, C2C12 and COS cells using the TRIzol reagent (GibcoBRL). The quality of the RNA was routinely assessed by size fractionation on formaldehyde-agarose gels followed by ethidium bromide staining. Northern blot hybridizations were performed by minor modifications of a previously described method.[15] Briefly, the RNA (10 μg) was size fractionated by formaldehyde-agarose gel electrophoresis and transferred to nylon membranes (Hybond-N, Amersham Buckinghamshire UK) by capillary transfer. Filters were baked at 80°C for 30 minutes followed by a UV cross-linking of the RNA at 0.4 J/cm^2. Filters were prehybridized for 30 minutes at 65°C using a RapidHyb Buffer (Amersham). MLC-2a cDNA containing plasmid[1] was digested with EcoR1 releasing a 400bp cDNA fragment. The fragment was isolated and purified with Qiagen (Hilden Germany), randomly labeled with α^{32}P dCTP (Dupont de Nemours NV, Brussels Belgium) and added to the prehybridization solution (activity:1.10^6 cpm/ml) for incubation during 2 h at 65°C. Stringent posthybridization wash conditions were used with 0.1 x SSC, 0.1% SDS at 65°C. Expression of the MLC-2a gene was detected by autoradiography for 16 hours with Kodak X-Omat (Rochester NY) film. As a control, for loading and transfer efficiency, the filters were routinely stained with methylene blue to visualize the 18 S and 28 S RNA bands. In addition hybridizations were performed using 18 S.

Creatine kinase assay

To assess the degree of differentiation of the C2C12 cells creatine kinase (CK) activity was determined. The method for determination of CK activity in differentiating C2C12 cells was the forward reaction, based on the *Tanzer* and *Gilvarg* method.[16] This method measures the change of extinction (rate of NADH removal) at 340nm in C2C12 lysate. This reaction depends on CK activity in the presence of added pyruvate kinase and lactate dehydrogenase.

Data analysis

Luciferase/Galactosidase (Luc/Gal) ratio was calculated to correct for differences in transfection efficiency within one experiment. To correct for variation in

consecutive experiments the Luc/Gal ratios are expressed relative to the Luc/Gal ratio of the 0.2 Kb MLC-2a promoter fragment, which was arbitrarily set at 1.0. For statistical analysis the non parametric Mann-Whithney U test and Kruskal Wallis test were applied. P-values of <0.05 were considered to be statistically significant.

Results

5' flanking sequence analysis

The 5' flanking region of the MLC-2a gene is 2045 bp in length. Upstream of the 5' flanking region, exon 11 the mouse glucokinase gene is located (chapter 6).[17] Several conserved DNA sites were recognized by the Sigscan software program. Sequence comparison with published cardiac specific elements[18-20] (for review chapter 2[21,22]), revealed the presence of additional putative *cis* elements. The localization of a selection of conserved transcription factor binding sites identified is shown in figure 1 and 2. The selection is based on *cis* elements and *trans* acting factors that have been shown to be important for contractile gene expression and cardiac development. No canonical TATA box was found, instead at position -28 an AT rich element is present (CATAA). The AT rich element could potentially interact with the TATA box binding protein to initiate transcription. Surrounding the transcription start site (-3,+5), a conserved Initiator (Inr)[23] site was identified (CTCAGTCA). The first 450 bp contain several *cis* elements that have been shown to positively regulate transcription in contractile genes, like GATA 4, Serum Response Factor (SRF), Transcription Enhancer Factor (TEF-1), CCAAT binding factor, Stimulating protein 1 (SP-1) and potentially the E-box could interact with cardiac bHLH proteins. Sigscan recognized seven conserved E-box elements within the 2045 bp flanking region, including one conserved upstream stimulating factor (USF) site at position -707 (Figure 2). Recent studies indicate an important role for hormones and morphogens like retinoic acid (RA), in cardiogenesis. Therefore we searched for DNA elements that can interact with members of the nuclear receptor superfamily. One conserved palindromic hormone receptor element (HRE -1622) site was found. A total of four retinoic acid responsive elements were recognized (RARE, -310, -871 in sense orientation and -926, -946 in antisense orientation). Transcription factors of the myocyte enhancer factor (MEF-2) family bind to A/T rich sites, two of which are localized at positions -1009 and -1152. The TGAAGAG sequence (-545) is a consensus binding site for the homeodomain containing protein, Nkx 2.5 *in vitro*.[20]

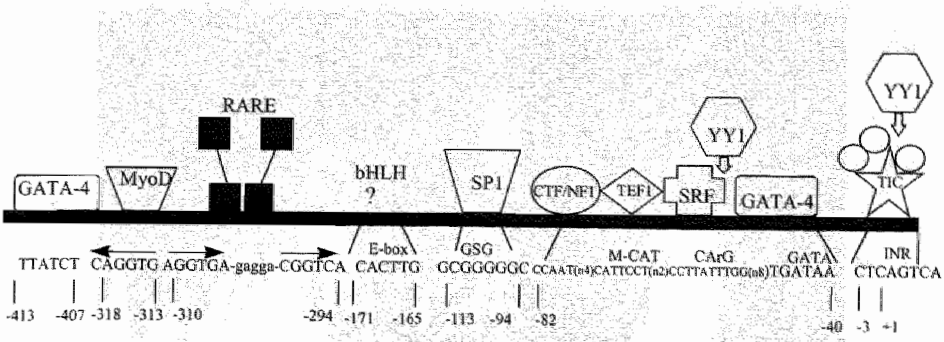


Figure 1. Putative *cis*-acting elements and potential *trans* acting factors within the 0.45 Kb MLC-2a promoter fragment. The position and sequence of the *cis* elements is shown below the bar. The putative *trans* acting factors are presented above the bar. Up to now it is unclear how many proteins are involved in the initiation of transcription, therefore in figure 1 at the conserved transcription initiation site (Inr) a transcription initiation complex (TIC) is indicated. This complex contains the RNA polymerase enzyme.

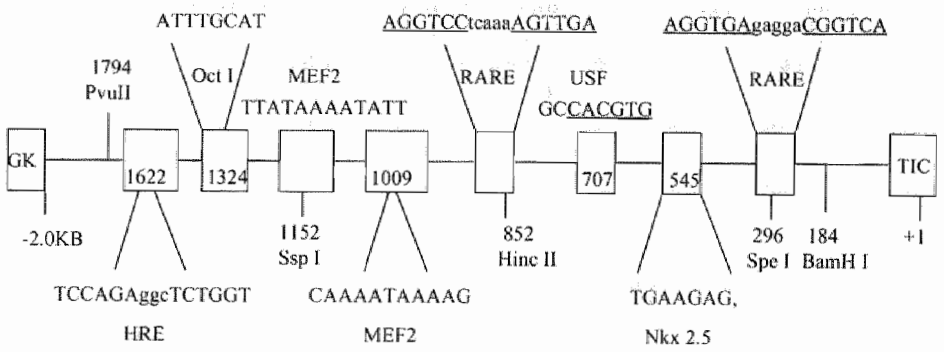


Figure 2. Putative *cis*-acting elements and potential *trans* acting factors on the 2.0 Kb MLC-2a promoter fragment. The position and sequence of the *cis* elements is shown and either the name of the *cis* element or the potentially *trans* acting factor is indicated.

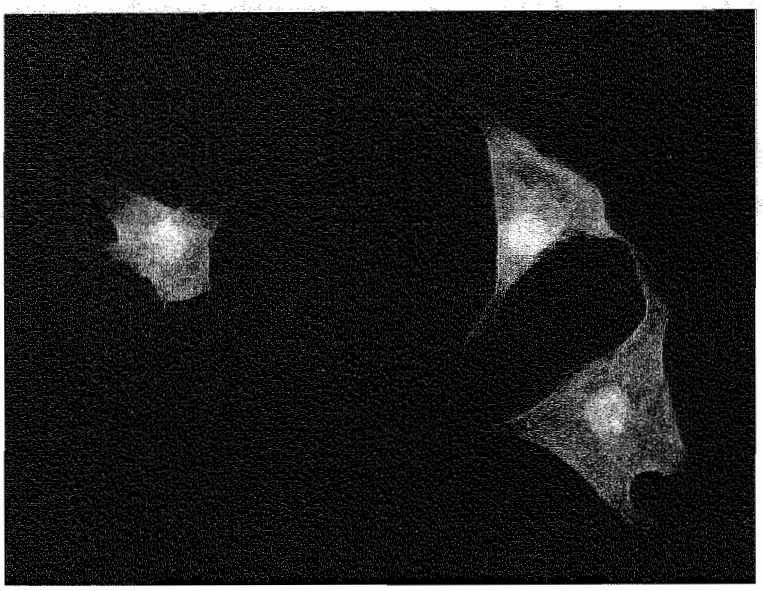


Figure 3A. Cell culture validation A) Purity of atrial myocyte cell culture. Immunostaining of cultured atrial cardiomyocytes with MLC-2a and fluorescein isothiocyanate (FITC) labeled donkey anti-rabbit immunoglobin in the presence of 4',6-diamidine-2®-phenylindole dihydrochloride (DAPI) for nuclear staining (magnification 40x).

Cell culture validation

The method for the purification of rat neonatal cardiomyocytes was adapted to isolate atrial myocytes. The purity of the cultured cardiomyocytes was checked by immunocytochemistry using the polyclonal MLC-2a antibody. The purity appeared to be > 95% in two separate isolations and staining after 3 days in culture (figure 3A).
Harvesting atrial myocytes is a time consuming procedure with a modest yield of myocytes allowing limited experiments (approximately $2x10^5$ myocytes from one neonatal rat heart). Therefore we analyzed the possibility to use genetically modified AT1 and AT2 atrial-like cardiomyocytes. To check for expression of the MLC-2a gene in AT1 and AT2 cells Northern blotting was performed (figure 3B). Both cell types express the MLC-2a gene and therefore provide an alternative cell line for transfection assays and transcription regulation studies. The differentiation of C2C12 cells from myoblasts into myotubes was confirmed by using creatine kinase as a marker, showing a threefold increase of CK activity for the myotube stage (figure 3A,B,C).

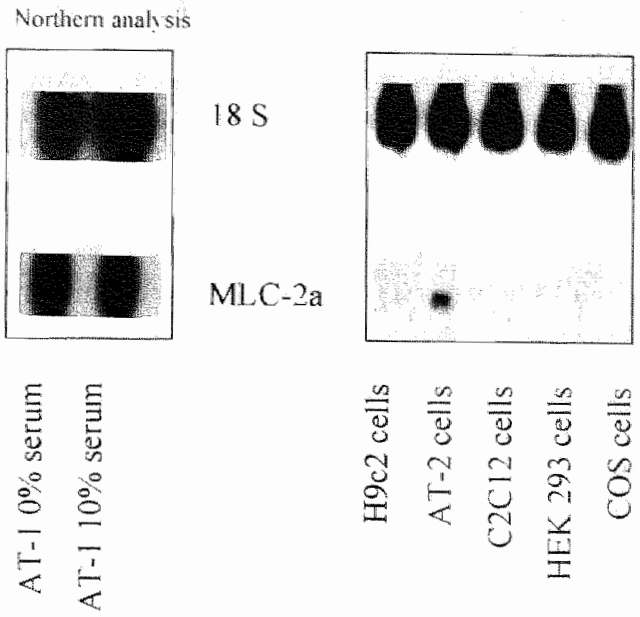


Figure 3B. Northern blot: To check whether the MLC-2a gene is expressed in AT1 and AT2 cells, Northern blot analysis was performed with total RNA, using murine MLC-2a cDNA as a probe. Note the presence of MLC-2a mRNA in AT1, AT2 cells and the absence of mRNA in H9c2, C2C12 (myotubes), HEK 293, and COS-7 cells.

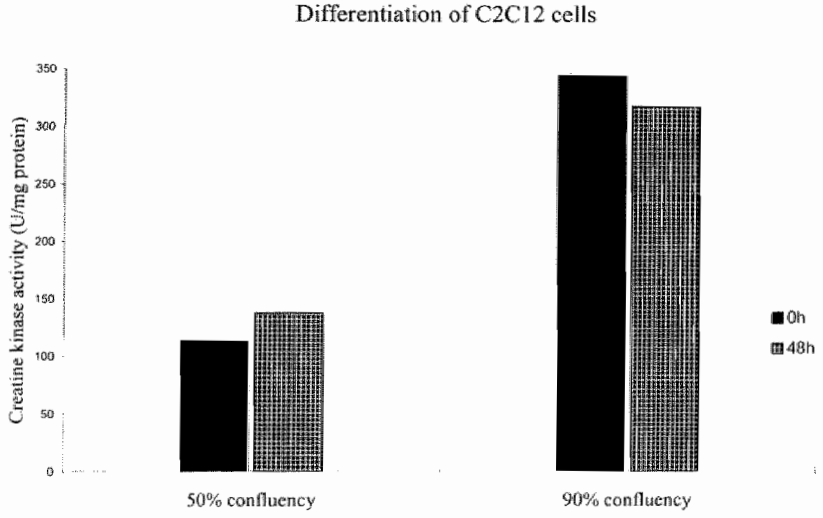


Figure 3C. Creatine kinase (CK) activity was measured in lysate from C2C12 myoblasts (50% confluency) and compared to C2C12 myotubes. (90% confluency) Ck activity levels were determined in cells at time 0, (before transfection) and after 48 hours following a standard calcium phosphate transfection experiments. In the myotube stage a threefold increase in CK activity was found indicating differentiation of C2C12 cells into myotubes.

Deletion fragment analysis

To identify the functionally important sites of the 5' flanking region of the MLC-2a gene, deletion fragments were subcloned into a promoterless luciferase vector. The constructs used for transient transfection assays are shown in figure 4A. The Luc/Gal ratios for RSV-luciferase and the 0.2 Kb fragment are shown in table 1. Calciumphosphate mediated transfections were not effective in AT1 and AT2 cells. However reproducible results were obtained after changing to liposome mediated transfections in AT1 cells, whereas levels in AT2 cells remained close to background (table 1). Therefore the results with AT2 cells are not included in the data analysis. Compared to RSV-luciferase, the highest level of luciferase expression after transfecting the 0.2 Kb MLC-2a promoter fragment is found in AT1 cells and ventricular cardiomyocytes. The expression level of the 0.2 Kb fragment in atrial cardiomyocytes is approximately 50% of that of ventricular cardiomyocytes. The 0.2 Kb construct is also active in C2C12 myoblasts, myotubes and COS cells albeit at low levels (8%, 5% and 9% of RSV-luciferase activity, respectively, table 1). These results indicate higher expression levels of the 0.2 Kb fragment in cardiac myocytes relative to non cardiac cells.

The deletion fragments of the MLC-2a promoter and 638ANF\luciferase were transfected into different cell types. In atrial and ventricular cardiomyocytes high levels of relative luciferase activity were calculated following transfection with the 0.45 promoter fragment relative to the 0.2 Kb fragment activity (figure 4B). Lowest activity was found for deletion fragment 0.9, and a modest increase was found by extending the promoter to 1.2 Kb. Again a modest repression is found in the 1.8 and 3.0 Kb promoter fragments. Overall these findings suggest strong silencing activity within the region between 0.45 and 0.9 Kb. Comparing the activity levels in atrial versus ventricular myocytes we found a strikingly similar pattern. None of the transfected constructs appeared to mediate atrial specific gene expression in rat neonatal myocyte cultures. Relative to the 0.2 Kb MLC-2a core promoter fragment, the 638ANF\luciferase construct was 3 fold more active in ventricular cardiomyocytes compared to atrial cardiomyocytes. The transfection results with AT1 cells were comparable to those in atrial and ventricular rat neonatal cardiomyocytes (figure 4C). Extending the regulatory sequence beyond 0.45 Kb markedly diminished the activity level to approximately 10% with the 0.9 Kb and 30% with the 1.2 Kb relative to the 0.2 Kb fragment.

To check whether the promoter mediates cardiac specific gene expression, skeletal myocyte cell line C2C12 was also transfected with MLC-2a promoter deletion fragments. Cells remain undifferentiated (myoblasts) at culture conditions of < 50% confluency and differentiate into myotubes at 90% confluency. In addition COS cells were studied as a representative non-muscle cell line. The transfection results are shown in figure 4D. The 0.2 Kb fragment displays low level of activity in C2C12 and COS cells. In C2C12 myoblasts the 0.2 Kb is slightly

Table 1. Activity of the 0.2 Kb MLC-2a promoter fragment relative to RSV-promoter activity in transient transfection experiments in different cell types.

Relative luciferase activity	RSV	MLC-2a	MLC-2a/RSV
Atrial myocytes	15972 ± 8754	2209 ± 1102	0.14
Ventricular myocytes	16818 ± 7552	4512 ± 1567	0.27
AT1	600 ± 159	193 ± 34	0.32
AT2	60 ± 40	10 ± 7	0.16
C2C12 myoblast	2179 ± 327	180 ± 38	0.08
C2C12 myotube	1356 ± 70	45 ± 12	0.05
COS	3474 ± 1158	313 ± 199	0.09

Relative activities of RSV and 0.2 Kb MLC-2a promoter fragment are presented as luciferase/galactosidase ratios to correct for differences in transfection efficiency. Data are mean ± sd of at least three independent experiments.

more active then in C2C12 myotubes (8% versus 5%, table 1), showing that initiation of sarcomere assembly is not associated with 0.2 Kb MLC-2a promoter activation. The longer promoter constructs (>0.2 Kb) are even less active in C2C12 myoblasts and myotubes and COS cells. Comparing the transfection results with the 0.45 Kb promoter fragment in cardiac myocytes (including AT1) to non cardiac cells, a statistically significant higher expression levels was found in the cardiac myocytes ($p < 0.05$, Mann-Whitney U rank sum test, figure 5).

Initiator mediated transcription regulation

To evaluate the importance of the conserved Inr site (-3,+5) for transcription initiation, Inr site mutation and deletion constructs were inserted in the 0.2 Kb promoter\luciferase construct by PCR. In mutation M1 and M2, two central bases were changed and in addition by partial (D1) and complete (D2) deletion the Inr site was altered. The wildtype and mutated sequences and constructs are shown in figure 6A. Mutation M2 changes the mRNA cap site (A to G). The mutated constructs were transfected into ventricular cardiomyocytes to assess the effect on luciferase expression levels. The level of activity of the mutated Inr constructs M1, M2, D1 and D2, were compared to the 0.2 Kb promoter fragment. M1 had no effect on transcription levels, whereas M2 showed a statistically significant 50% reduction. Furthermore partial or complete deletion of Inr (D1 and D2) reduced reporter gene expression significantly ($p<0.05$ Kruskal Wallis test). In all mutated constructs the AT rich CATAA site (position -28) is intact. The importance of an intact Inr site for transcription initiation is evident, although he impact of the double base mutations (M1 and M2) is limited.

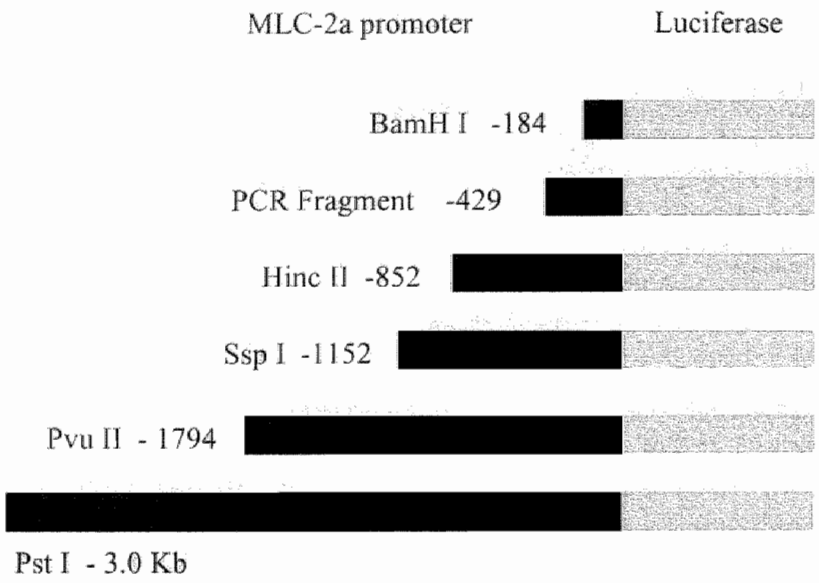


Figure 4A. A Deletion constructs of MLC-2a promoter: The position of the restriction site in the MLC-2a 5' flanking region used to make the constructs is indicated. The plasmids are referred to as 0.2 Kb (-184 BamH I), 0.45 (PCR fragment), 0.9 (-852 Hinc II), 1.2 (-1152 Ssp I), 1.8 (-1794 Pvu II), 3.0 (-3003 Pst I).

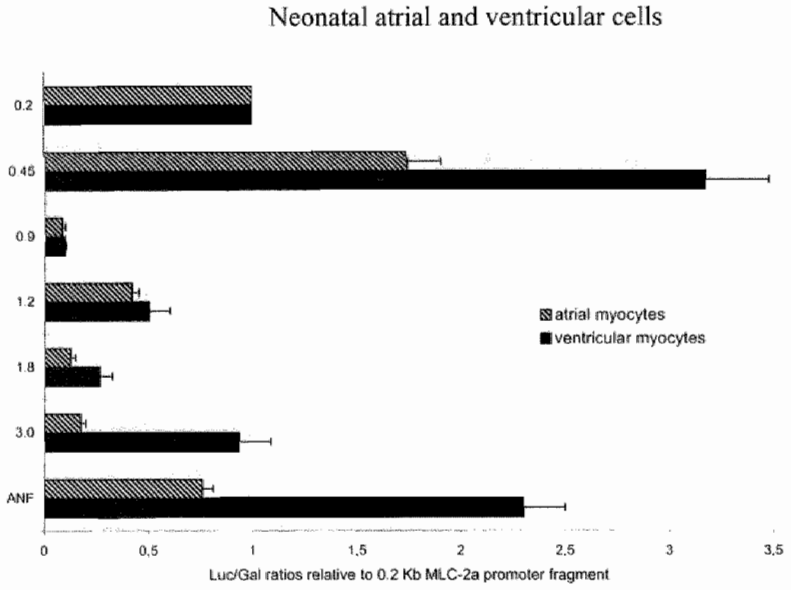


Figure 4B. Relative luciferase activity for the MLC-2a/luciferase constructs in atrial and ventricular cardiomyocytes in the presence of 10% serum. To correct for variation in consecutive experiments the Luc/Gal ratios are expressed relative to the Luc/Gal ratio of the 0.2 Kb MLC-2a promoter fragment, which was arbitrarily set at 1.0.

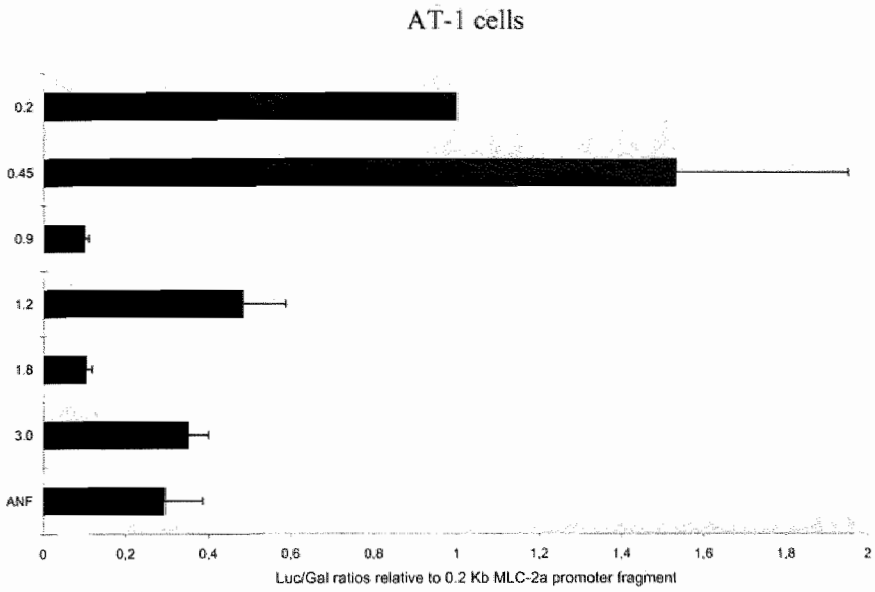


Figure 4C. Relative luciferase activity for the MLC-2a/luciferase constructs in AT1 cells.

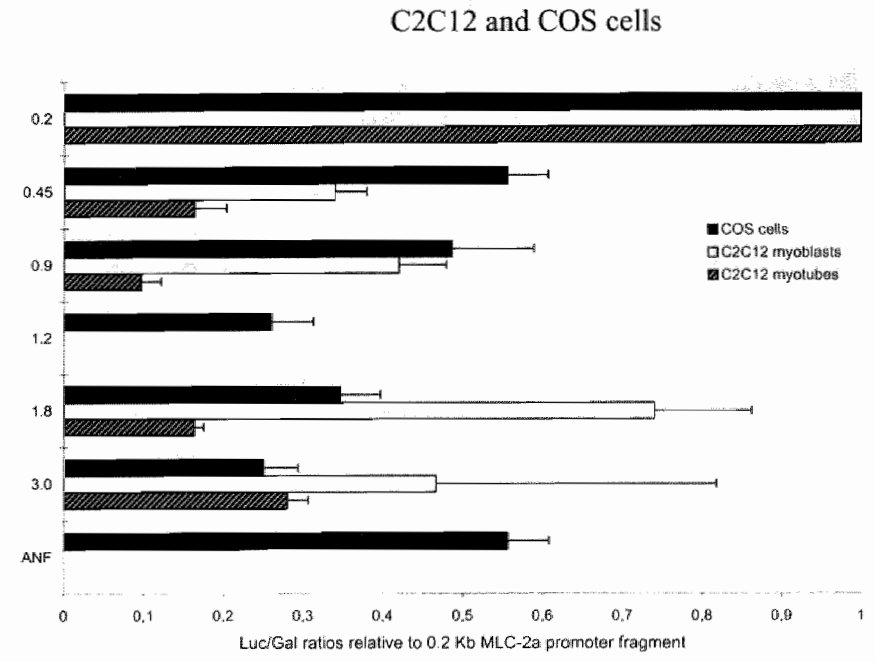


Figure 4D. Relative luciferase activity for the MLC-2a/luciferase constructs in COS cells, C2C12 myoblasts and C2C12 myotubes.

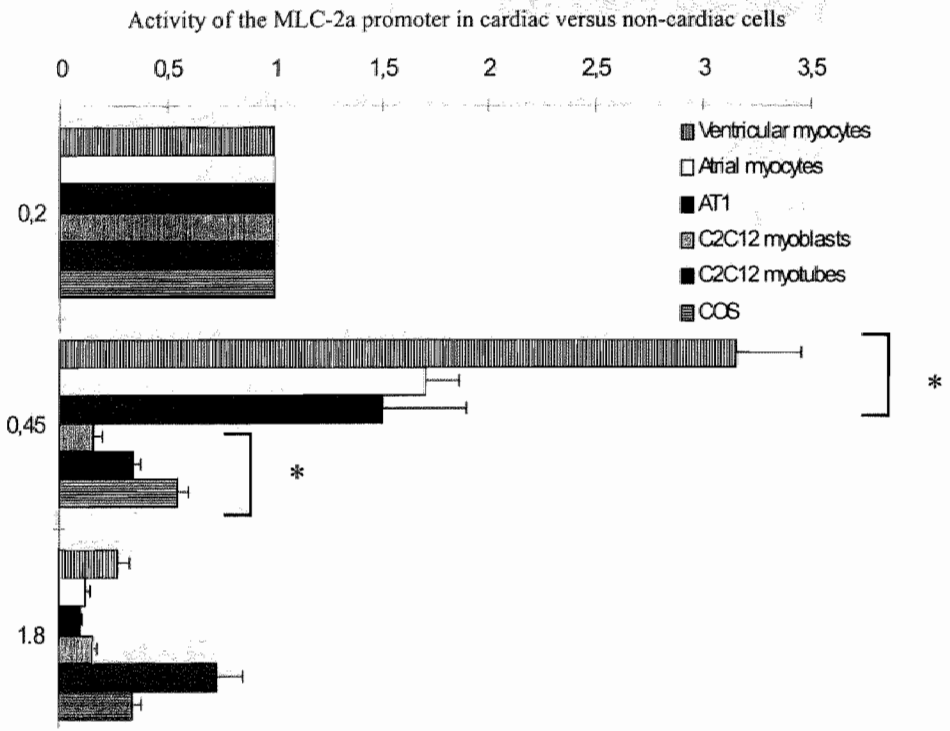


Figure 5. Comparison of transfection results in 0.2, 0.45 and 1.8 Kb promoter fragment in atrial, ventricular neonatal cardiomyocytes, AT1, COS and C2C12 cells. A statistical significant difference was found ($p<0.05$, Mann Whithney U test) for the 0.45 Kb fragment comparing cardiac (including AT1) with non cardiac cells.

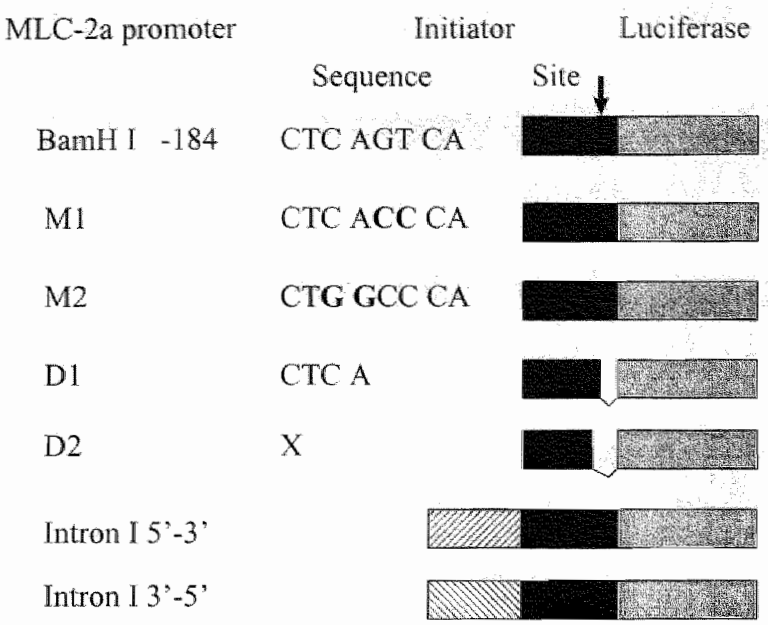


Figure 6A. Schematic representation of the mutation, deletion and insertion constructs made with the 0.2 Kb MLC-2a promoter fragment. The MLC-2a Inr wild type sequence is shown in the upper construct (BamH I), the mutations are indicated in bold (M1 and M2) and the deletions are shown. Intron 1 was PCR amplified and cloned upstream of the 0.2 Kb promoter fragment.

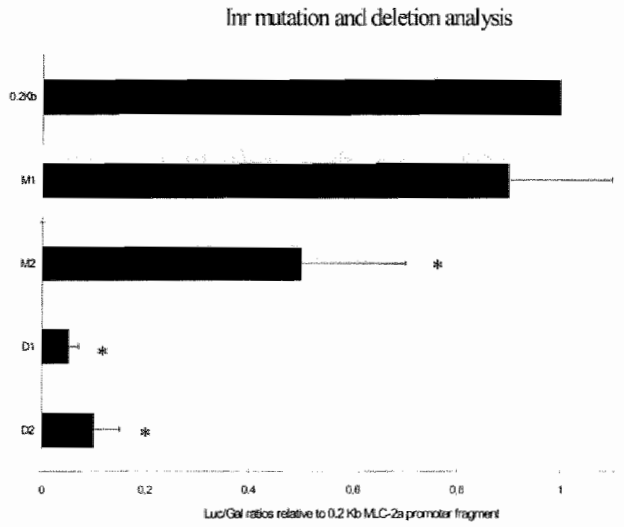


Figure 6B. Inr site deletions and mutations and promoter activity measured by transient transfection assays in ventricular cardiomyocytes. To correct for variation in consecutive experiments the Luc/Gal ratios are expressed relative to the Luc/Gal ratio of the 0.2 Kb MLC-2a promoter fragment, which was arbitrarily set at 1.0. M2, D1, D2 showed a significant downregulation of luciferase activity ($p<0.05$, Kruskal Wallis).

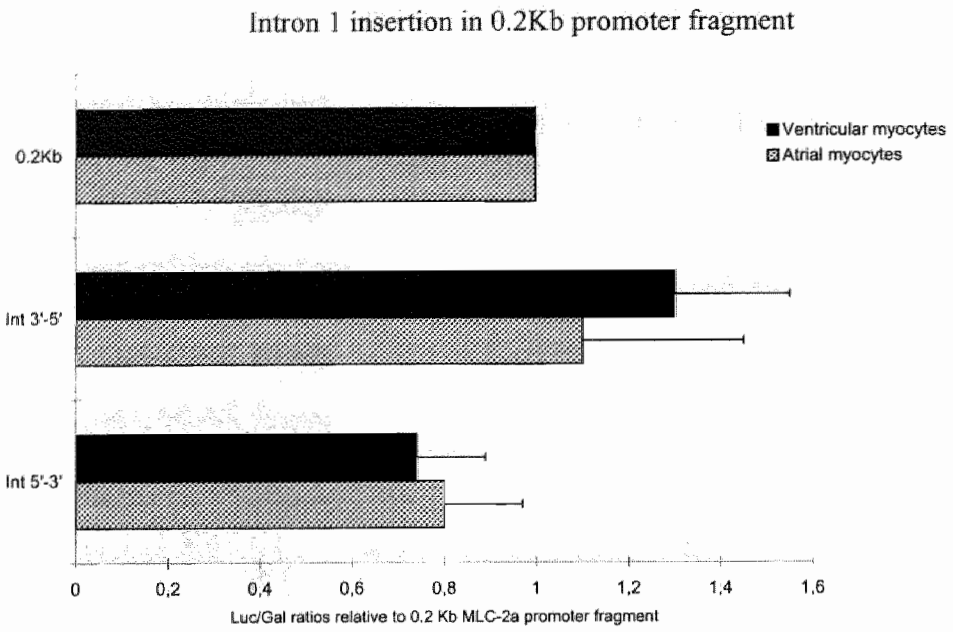


Figure 6C. Effect of addition of intron 1 on promoter activity. Intron 1 was inserted upstream of the 0.2 Kb promoter in both orientations. Neonatal ventricular myocytes were transiently transfected with the constructs indicated (A).

Intron 1 regulatory sites

Various groups reported the presence of regulation sequences within the first intron.[24,25] To test whether intron 1 of the MLC-2a gene contains important regulatory elements, the first intron was PCR amplified and cloned upstream of the 0.2 Kb promoter fragment in both orientations (Figure 6A). The constructs were transfected in both atrial and ventricular myocytes. There was no change in expression levels due to the presence of intron 1, in ventricular or atrial myocytes.

Discussion

In this study transient transfection with promoter deletion fragments was performed to identify potential regulatory DNA regions responsible for atrial specificity. The short MLC-2a promoter fragments (0.2 and 0.45 Kb) are active in atrial cardiomyocytes but show higher levels of activity in ventricular cardiomyocytes.

The 0.45 Kb fragment of the 5' flanking region of the MLC-2a gene confers cardiac specific gene expression in culture conditions. Therefore this part of the 5' flanking region of the MLC-2a gene must contain important *cis* element binding to upregulating transcription factors. The proximal (0.2 Kb deletion fragment) putative essential promoter is active, albeit at low levels, in C2C12 myoblasts and COS cells. Strong silencing elements are located between 0.45 and 0.9 Kb of the promoter fragment. No atrial specific enhancer fragment was identified through transient transfection assays in cultured atrial and ventricular cardiomyocytes under the conditions applied.

Positive transcription regulating elements

The cardiomyocytes were maintained on serum supplemented medium after transfection. The main reason for this treatment is the healthy appearance of serum treated cardiomyocytes, compared to cells cultured in serum free conditions. The addition of serum can induce hypertrophy in cultured cardiomyocytes. Therefore the addition of serum could potentially alter the MLC-2a promoter activity. Indeed the relatively high activity level of the ANF promoter could indicate a hypertrophic response of the cardiomyocytes.

The 0.45 Kb is significantly stronger in cardiomyocytes compared with non cardiac cells. This promoter fragment has a higher activity in ventricular than atrial cells. The consensus *cis* elements known to be relevant for contractile gene expression are shown in figure 1. At position -40 a conserved GATA element is present. Up to now six different GATA-binding proteins interacting with the conserved sequence of (A/T)GATA(A/G) have been cloned. The GATA-4 is an interesting factor as the expression of this zinc finger protein turned out to be predominantly in the heart. GATA-4 was not detected in skeletal muscle.[26,27]

TEF-1, interacting with the M-CAT box is another candidate transcription factor for the MLC-2a promoter. TEF-1 is ubiquitously expressed but the highest level can be found in cardiac and skeletal muscle. TEF-1 is also involved in the signaling pathway of α-adrenergic stimulation.[28,29] Factors like Sp-1 interacting with the GSG box and SRF binding to the CArG element are ubiquitous factors. Potentially homeobox genes could bind to the CArG box as well.[30] These elements are located within the first 203 bp and although this MLC-2a promoter fragment is more active in cardiac versus non cardiac cells, the activity is not restricted to cardiomyocytes. In the region spanning from -0.45 to -0.2 Kb, consensus sites are found at position -310 (RARE) -318 (E-box) and -413 (GATA). In addition to members of the GATA family, other transcription factors present in cardiomyocytes can interact with these sites. The RARE can bind various combinations of members of the nuclear receptor superfamily. These receptors can influence transcription.[31,32] The E-box can bind bHelix-Loop-Helix (bHLH) proteins. Although proteins belonging to the bHLH have been identified in the heart their

role in transcription regulation is still unclear.[33,34] It is crucial to realize that other factors, not yet identified, may be involved in transcription regulation of the MLC-2a promoter.

Negative transcription regulating elements

A marked downregulation in relative luciferase activity was noticed using the longer deletion fragments (0.9 Kb) compared to the short fragments (0.45 Kb), which is suggestive for the presence of a silencer in this region. This silencer sequence appears to be active in both atrial and ventricular cardiomyocytes. The regulatory DNA of the MLC-2a gene contains several interesting putative binding sites between position 0.45 and 0.9 Kb, including Nkx 2.5 (-545), USF (-707) and RARE (-871) consensus sites.

The importance of hormones and morphogens like RA for cardiac gene expression is well established.[35-37] Both up- and downregulation of transcription through RARE has been demonstrated.[31,38] The role of RA, its receptors and DNA binding element during embryonic development expression was supported by Studer et al.[38] They showed repression of the homeobox gene, Hoxb-1 during rhombomere formation in the hindbrain. The repression was defective in transgenics harboring mutated RARE in the silencer of the Hoxb-1 promoter. In addition a role for RARE was found in the promoter of the chicken lysozyme gene. The silencer of the lysozyme gene is inactive in lysozyme expressing macrophages, but active in all other cell types studied. Cipitelli et al[39] studied the effect of RA on transcriptional regulation of the IFN-γ promoter. RA treatment downregulates the IFN-γ promoter. This inhibitory effect was reproduced by RAR-α overexpression and also required an adjacent intact E-box *cis* element, previously characterized as an USF binding element.

USF belongs to the family of the basic Helix-Loop-Helix proteins. The USF protein has been shown to interact with several conserved E-box elements of the ventricular MLC-2 promoter. The USF-DNA interaction can lead to both up- or downregulation of transcription, depending on the *cis*-element.[34] Activity of the ventricular MLC-2 promoter (250 bp) is ventricular myocyte specific. There is evidence suggesting that USF is part of the DNA-protein complex mediating this ventricular myocyte specific promoter activity. The MLC-2a promoter has a sequence composition, where a RARE is located at close proximity to a consensus USF binding site (E-box), like IFN-γ. The relation between the RARE and USF binding site on the MLC-2a promoter is particularly interesting and needs further evaluation.

Homeodomain containing proteins can play a role in transcriptional silencing. The homeodomain protein Oct-1 reduces transcription through AT rich *cis* elements on the promoter of the human thyrotropin beta gene.[40] However in the 0.45-0.9 spanning region no AT rich elements were encountered. In addition to

AT rich elements alternative sites like the putative Nkx 2.5 element at position -545 can interact with homeodomain containing proteins. Whether this *cis* element is important remains to be determined.

Initiator cis element (Inr)

The genomic structure of the various myosin light chain isoforms in different species is highly conserved (chapter 6). Transcription initiation is presumed to be TATA box dependent in the murine (m) mMLC-1a, mMLC-1v, mMLC-1f and mMLC-2v genes although the canonical TATAA sequence is not found in any of these genes. Instead an AT-rich element or TATA box like element is found at position between -32 and -25, in the four MLC genes (mMLC-1f/3f promoter: TATATA, mMLC-1v: TATAT, mMLC-1a, AATAAAA[41], mMLC-2v TATTTAT[42]). However in the 5' flanking region of the MLC-2a gene no conserved AT rich element is found close to the transcriptional start site, except CATAA. The initial C of this DNA sequence makes a function as TATA box unlikely.[43] Interestingly a conserved *cis* element was found designated as initiator (Inr) by Smale et al.[23] The Inr element was previously identified on the deoxynucleotidyl transferase gene. The conserved Inr is absent in the 5' flanking region of the mMLC 1f/3f, 1v, 2v, 1a genes. *In vitro* transcription initiation appeared to require an Sp1-dependent activating site upstream (GSG-box).[23] The consensus sequence of the Inr site is CTCANTCT (CTCANTCA in MLC-2a, figures 1,3). Using deletion and mutation analysis, the importance of the Inr site was shown. Partial or complete deletion of the Inr site completely attenuated luciferase transcription. The effect of mutations M1 and M2 was less dramatic. Transcription was normal in the presence of M1 and reduced to 50% in M2. Smale et al. reported complete attenuation of *in vitro* testtube transcription after the introduction of mutation M2 in an artificial promoter construct. This could indicate the presence of adjacent important *cis* elements in the setting of the MLC-2a promoter. The change of the mRNA cap site introduced in M2 is apparently not detrimental for the transcript. Several transcription factors have been reported to interact specifically with the Inr element including USF[44], YY1[45], and replication protein A.[46] Recent studies by Manzano-Winkler et al.[43] showed the requirement of the transcription factor II-D for promoters with and without TATA box. However TATA-less promoters containing exclusively an Inr element need an additional protein for transcription initiation called transcription factor (TF) II-I. These studies were performed using the T cell receptor variable region-derived promoter. TF II-I has striking similarities to USF with respect to binding characteristics and immunological recognition.[44] An increasing number of genes appears to be Inr dependent including tissue specific genes like MLC-2a (this study) and the endothelial cell specific vascular endothelial growth factor receptor, the Flk-1 gene.[47] The functional importance of the Inr site for transcription initiation of the MLC-2a gene

was demonstrated by MLC-2a Inr mutations in transient transfection assays in cultured ventricular cardiomyocytes.

Introns and transcriptional regulation

The general function of introns is unclear. Bacterial genes have no introns. Eukaryotic genes like insulin, however appear to have sometimes shorter and even loss of introns, in more complex organisms.[48,49] In contrast the size of the introns in the MLC-2a gene is increasing throughout evolution (chapter 6). One important advantage of the presence of introns is the possibility of alternative splicing a mechanism leading to various protein products from one primary transcript. Several cell type specific enhancers have been shown to be located within the first intron.[24] As the orientation of an enhancer is not crucial for its function the first intron was placed upstream of the essential promoter of the MLC-2a gene, in both orientations. Addition of the intron had no significant effect on essential promoter (0.2 Kb) activity in atrial and ventricular cells. Therefore it is unlikely that regulatory sequences important for transcription in cardiomyocytes are located within this region.

MLC-2a transcription as a marker in cardiogenesis

The MLC-2a gene is initially expressed in the entire heart tube from day 8 of murine heart development. The expression of the gene becomes restricted to the atrial compartment in a short time frame around day 12. In contrast the expression of the ventricular isoform (MLC-2v) is limited to the ventricular compartment throughout development. In knock out experiments of one of the retinoic acid receptors (retinoid X receptor-α) this pattern is disturbed. Homozygous (-/-) knockout embryos, die of pump failure at day 13-14 and histological analysis revealed major cardiac defects. There is a thin ventricular wall, less trabeculations and a persistent ventricular septal defect. On the RNA level a persistent expression of the MLC-2a isoform was found. Therefore the regulatory sequence of the MLC-2a gene must contain important clues, to explain the molecular mechanism leading to persistence of MLC-2a expression in the RXR-α knockout ventricular chamber. Similar results were obtained in the homeobox (Nkx 2.5) knock-out, where the absence of MLC-2v isoform expression was seen in the ventricular compartment of (-/-) embryos. These embryos die around day 11 and a persistent heart tube, lacking ventricular growth is found. This inhibition of ventricular growth is the most important histological finding.[50] In the heart tube of the (-/-) embryos MLC-2a is present. Thus even in the absence of Nkx 2.5, the MLC-2a promoter is activated.

Conclusions

Through *in vitro* characterization of the regulatory sequence of the MLC-2a gene, a cardiac specific fragment of 0.45 Kb in length was identified. The promoter fragment spanning from -203 to -450 bp enhances transcription, whereas as the -450 to -900, acts as a silencer. Several putative *cis* elements involved in transcription regulation of the MLC-2a gene have been recognized. A conserved *cis*-element functions as the transcription initiator site. However no functional atrial specific regulatory element was found in transient transfection assays comparing reporter gene activity levels in atrial versus ventricular cardiomyocytes.

Literature

1. Kubalak SW, Miller-Hance W, O'Brien T, Dyson E, Chien K: Chamber-specification of atrial myosin light chain-2 expression precedes septation during mouse cardiogenesis. *J Biol Chem* 1994;269:16961-16970
2. O'Brien TX, Lee KJ, Chien KR: Positional specification of ventricular myosin light chain 2 expression in the primitive murine heart tube. *Proc Natl Acad Sci U S A* 1993;90:5157-5161
3. Miller-Hance WC, LaCorbiere M, Fuller SJ, Evans SM, Lyons G, Schmidt C, Robbins J, Chien KR: In vitro chamber specification during embryonic stem cell cardiogenesis. Expression of the ventricular myosin light chain-2 gene is independent of heart tube formation. *J Biol Chem* 1993;268:25244-25252
4. Lin Q, Schwarz J, Bucana C, Olson EN: Control of mouse cardiac morphogenesis and myogenesis by transcription factor MEF2C. *Science* 1997;276:1404-1408
5. Rump R, Buhlmann C, Borchers T, Spener F: Differentiation-dependent expression of heart type fatty acid-binding protein in C2C12 muscle cells. *Eur J Cell Biol* 1996;69:135-142
6. Kubalak S, Doevendans PA, Rockman H, Hunter JJ, Tanaka N, Ross J, Chien KR: Molecular analysis of cardiac muscle diseases based on mouse genetics, in Adolph KW (ed): *Human molecular genetics.* Orlando, Academic Press, 1996, pp 470-487
7. Simpson P, Shoshana S: Differentiation of rat myocytes in single cultures with and without proliferating non-myocardial cells. *Circ Res* 1982;50:101-116
8. Chen C, Okayama H: High-efficiency transformation of mammalian cells by plasmid DNA. *Mol Cell Biol* 1987;7:2745-2752
9. Field L: Atrial Natriuretic Factor-SV40 T antigen transgenes produce tumors and cardiac arrhythmias in mice. *Science* 1988;239:1029-1033
10. Borisov AB, Claycomb WC: Proliferative potential and differentiated characteristics of cultured cardiac muscle cells expressing the SV40 T oncogene. *Ann N Y Acad Sci* 1995;752:80-91

11. de Wet JR, Wood KV, DeLuca M, Helinski D, Subramani S: Firefly luciferase gene: structure and expression in mammalian cells. *Mol Cell Biol* 1986;7:725-737
12. MacGregor GR, Nolan GP, Fiering S, Roederer M, Herzenberg LA: Use of *E. coli lacZ* (β-galactosidase) as a reporter gene, in Murray EJ (ed): *Methods in molecular biology.* Clifton, NY, The Humana Press, 1991, pp 217-235
13. Thorburn A, Thorburn J, Chen SY, Powers S, Shubeita HE, Feramisco JR, Chien KR: HRas-dependent pathways can activate morphological and genetic markers of cardiac muscle cell hypertrophy. *J Biol Chem* 1993;268:2244-2249
14. Nordeen SK: Luciferase reporter gene vectors for analysis of promoter and enhancers. *Biotech* 1988;6:454-457
15. Trahair T, Yeoh T, Cartmill T, Keogh A, Spratt P, Chang V, dos Remedios CG, Gunning P: Myosin light chain gene expression associated with disease states of the human heart. *J Mol Cell Cardiol* 1993;25:577-585
16. Tanzer ML, Gilvarg C: Creatine and creatine kinase measurement. *J Biol Chem* 1959;234:3201-3204
17. Postic C, Niswender KD, Decaux JF, Parsa R, Shelton KD, Gouhot B, Pettepher CC, Granner DK, Girard J, Magnuson MA: Cloning and characterization of the mouse glucokinase gene locus and identification of distal liver-specific DNase I hypersensitive sites. *Genomics* 1995;29:740-750
18. Mangelsdorf DJ, Evans RM: The RXR heterodimers and orphan receptors. *Cell* 1995;83:841-850
19. Beato M, Herrlich P, Schutz G: Steroid hormone receptors: many actors in search of a plot. *Cell* 1995;83:851-857
20. Chen CY, Schwartz RJ: Identification of novel DNA binding targets and regulatory domains of a murine tinman homeodomain factor, nkx-2.5. *J Biol Chem* 1995;270:15628-15633
21. Doevendans PA, van Bilsen M: Transcription factors and the cardiac gene programme. *Int J Biochem & Cell Biol* 1996;28:387-403
22. Evans RM: The steroid and thyroid hormone receptor superfamily: transcriptional regulators of development and physiology. *Science* 1988;240:889-895
23. Smale ST, Schmidt MC, Berk AJ, Baltimore D: Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID. *Proc Natl Acad Sci U S A* 1990;87:4509-4513
24. Parmacek MS, Vora AJ, Shen T, Barr E, Jung F, Leiden JM: Identification and characterization of a cardiac-specific transcriptional regulatory element in the slow/cardiac troponin C gene. *Mol Cell Biol* 1992;12:1967-1976
25. Berx G, Staes K, van Hengel J, Molemans F, Bussemakers MJ, van Bokhoven A, van Roy F: Cloning and characterization of the human invasion suppressor gene E-cadherin (CDH1). *Genomics* 1995;26:281-289
26. Laverriere AC, MacNeill C, Mueller C, Poelmann RE, Burch JB, Evans T: GATA-4/5/6, a subfamily of three transcription factors transcribed in developing heart and gut. *J* Biol Chem 1994;269:23177-23184

27. Heikinheimo M, Scandrett JM, Wilson DB: Localization of transcription factor GATA-4 to regions of the mouse embryo involved in cardiac development. *Dev Biol* 1994;164:361-373
28. Kariya K, Karns LR, Simpson PC: 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* 1994;269:3775-3782
29. Karns LR, Kariya K, Simpson PC: M-CAT, CArG, and Sp1 elements are required for alpha 1-adrenergic induction of the skeletal alpha-actin promoter during cardiac myocyte hypertrophy. Transcriptional enhancer factor-1 and protein kinase C as conserved transducers of the fetal program in cardiac growth. *J Biol Chem* 1995;270:410-417
30. Lints TJ, Parsons LM, Hartley L, Lyons I, Harvey RP: Nkx-2.5: a novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants. *development* 1993;119:419-431
31. Mangelsdorf DJ, Umesono K, Kliewer SA, Borgmeyer U, Ong ES, Evans RM: A direct repeat in the cellular retinol binding protein type II gene confers differential regulation by RXR and RAR. *Cell* 1991;66:555-561
32. Durand B, Saunders M, Leroy P, Leid M, Chambon P: All-trans and 9-cis retinoic acid induction of CRABPII transcription is mediated by RAR-RXR heterodimers bound to DR1 and DR2 repeated motifs. *Cell* 1992;71:73-85
33. Srivastava D, Cserjesi P, Olson EN: A subclass of bHLH proteins required for cardiac morphogenesis. *Science* 1995;270:1995-1999
34. Navankasattusas S, Sawadogo M, van Bilsen M, Dang CV, Chien KR: The basic helix-loop-helix protein upstream stimulating factor regulates the cardiac ventricular myosin light-chain 2 gene via independent cis regulatory elements. *Mol Cell Biol* 1994;14:7331-7339
35. Dyson E, Sucov HM, Kubalak SW, Schmid Schonbein GW, DeLano FA, Evans RM, Ross J, Jr., Chien KR: Atrial-like phenotype is associated with embryonic ventricular failure in retinoid X receptor alpha -/- mice. *Proc Natl Acad Sci U S A* 1995;92:7386-7390
36. Olson EN, Srivastava D: Molecular pathways controlling heart development. *Science* 1996;272:671-676
37. Sucov HM, Dyson E, Gumeringer CL, Price J, Chien KR, Evans RM: RXRa mutant mice establish a genetic basis for vitamin A signaling in heart morphogenesis. *Genes Dev* 1994;8:1007-1018
38. Studer M, Popperl H, Marshall H, Kuroiwa A, Krumlauf R: Role of a conserved retinoic acid response element in rhombomere restriction of Hoxb-1. *Science* 1994;265:1728-1732
39. Cippitelli M, Ye J, Viggiano V, Sica A, Ghosh P, Gulino A, Santoni A, Young HA: Retinoic acid-induced transcriptional modulation of the human interferon-gamma promoter. *J Biol Chem* 1996;271:26783-26793

40. Kim MK, Lesoon Wood LA, Weintraub BD, Chung JH: A soluble transcription factor, Oct-1, is also found in the insoluble nuclear matrix and possesses silencing activity in its alanine-rich domain. *Mol Cell Biol* 1996;16:4366-4377
41. van Bilsen M, Chien KR: Growth and hypertrophy of the heart: towards an understanding of cardiac specific and inducible gene expression. *Cardiovasc Res* 1993;27:1140-1149
42. Henderson S, Spencer M, Sen A, Kumar C, Siddiqui M, Chien K: Structure, organization, and expression of the rat cardiac myosin light chain-2 gene. Identification of a 250-base pair fragment which confers cardiac-specific expression. *J Biol Chem* 1989;264:18142-18148
43. Manzano Winkler B, Novina CD, Roy AL: TFII is required for transcription of the naturally TATA-less but initiator-containing Vbeta promoter. *J Biol Chem* 1996; 271:12076-12081
44. Roy AL, Meisterernst M, Pognonec P, Roeder RG: Cooperative interaction of an initiator binding transcription initiator factor and the helix-loop-helix activator USF. *Nature* 1991;354:245-248
45. Seto E, Shi Y, Shenk T: YY1 is an intiator sequence binding protein that directs and activates transcription *in vitro*. *Nature* 1991;354:241-245
46. Tang CM, Tomkinson AE, Lane WS, Wold MS, Seto E: Replication protein A is a component of a complex that binds the human metallothionein IIA gene transcription start site. *J Biol Chem* 1996;271:21637-21644
47. Ronicke V, Risau W, Breier G: Characterization of the murine FLK-1 promoter. *Circ Res* 1996;79:277-285
48. Bell GI, Pictet RL, Rutter WJ, Cordell B, Tischer E, Goodman HM: Sequence of the human insulin gene. *Nature* 1980;284:26-32
49. Perler F, Efstratiadis A, Lomedico P, Gilbert W, Kolodner R, Dodgson J: The evolution of genes: the chicken preproinsulin gene. *Cell* 1980;20:555-566
50. Lyons I, Parsons LM, Hartley L, Li R, Andrews JE, Robb L, Harvey RP: Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene Nkx2-5. *Genes Dev* 1995;9:1654-1666

Chapter

The regulatory atrial myosin light chain gene is re-expressed in ventricular myocytes under hypertrophic conditions

Pieter A. Doevendans, Ronald Bronsaer, Jan M. van Dantzig, Kees Prenger, Christian Grohe, Pilar Ruiz Lozano, Hein J.J. Wellens, Marc van Bilsen

Abstract

Background: During development the atrial myosin light chain-2 gene is expressed in the whole heart tube. From day 12 the gene is silenced in ventricular myocytes. However in hypertrophic cardiomyocytes fetal genes can be reactivated. Therefore the re-expression of the endogenous atrial MLC-2 gene was evaluated under hypertrophy inducing conditions both in cultured rat cardiomyocytes and in tissue from patients with idiopathic dilated cardiomyopathy. In addition the effect of neurohormonal treatment on transfected MLC-2a promoter constructs was determined in cultured ventricular myocytes.

Methods: In primary cultures of rat neonatal ventricular cardiomyocytes the influence of serum conditions, phenylephrine, endothelin-1 and β-estradiol on MLC-2a mRNA expression was studied by Northern Blot and rt-PCR. In addition cardiomyocytes were transfected with plasmids containing deletion fragments of the MLC-2a promoter linked to the reporter gene luciferase. Following transfection the cardiomyocytes were treated with hypertrophic stimuli and β-estradiol and after cell lysis luciferase activity was determined.

Results: The endogenous MLC-2a gene is expressed in cultured cardiomyocytes at a low level in the absence of serum and expression increases with higher serum concentrations. The exogenous MLC-2a constructs containing >1.8 Kb of the promoter are activated by phenylephrine and endothelin-1. In addition the long fragments >1.8 Kb are activated by β-estradiol treatment. In contrast no induction of the short promoter constructs (<1.8 Kb) was observed after neurohormonal stimulation. In patients with dilated cardiomyopathy the MLC-2a gene is re-expressed in ventricular tissue as shown on the RNA and protein level.

Conclusions: The MLC-2a gene is re-expressed under hypertrophic stimuli in cultured rat neonatal cardiomyocytes and in human ventricular tissue from patients with dilated cardiomyopathy. These findings suggest that the MLC-2a gene is part of the fetal gene program activated during dedifferentiation induced by hypertrophic stimuli in ventricular cardiomyocytes.

Introduction

The mammalian heart adapts to increased hemodynamic load such as that induced by hypertension, valvular heart disease, and loss of cardiomyocytes (as with myocardial infarction or cardiomyopathy), by developing compensatory hypertrophy of the remaining cells.[1,2] The molecular response to hypertrophic stimuli of ventricular myocytes is characterized by dedifferentiation.[3] This molecular dedifferentiation involves re-expression of genes that were transiently expressed during embryonic development. This molecular change could indicate an attempt of cardiomyocytes to return to a stage, in which DNA duplication and mitosis were possible. Key players in mediating the hypertrophic response are the receptors for sensing the hypertrophic stimulus and intracellular signaling pathways, which ultimately give rise to selective activation of transcription factors, and finally expression of fetal genes, like atrial natriuretic factor and skeletal α-actin. The molecular response evoked depends on the stimulus applied. The re-expression of fetal genes is the final outcome of a cascade of molecular events.[2,4,5]

The murine regulatory myosin light chains provide chamber specific molecular markers during embryonic development, as the ventricular myosin light chain 2 (MLC-2v) is ventricular chamber specific at all stages. In contrast the expression of atrial myosin light chain 2 (MLC-2a) is initially found throughout the developing heart tube but becomes restricted to the atrial compartment from day 12 post coitum.[6] Therefore the MLC-2a gene is a fetal gene transiently expressed in ventricular myocytes.

The molecular response of ventricular myocytes to hypertrophic stimuli is species dependent. For instance the contractile apparatus of rodent cardiomyocytes is composed of other myosin heavy chain protein isoforms (predominantly α) compared to the human situation where β myosin is the predominant isoform. During hypertrophy in rodent the α isoform is replaced by β, while no significant change is observed in the human ventricular myocyte. In both species however the atrial natriuretic factor (ANF) gene is re-expressed.[7]

In this study we describe the behavior of the atrial MLC-2 isoform after application of various well-known hypertrophic stimuli and after estrogen treat-ment in rat ventricular neonatal myocytes. In addition transient transfection experiments were performed using deletion fragments of the MLC-2a regulatory DNA driving the luciferase reporter gene to identify inducible promoter fragments. Finally human tissue samples of patients undergoing partial ventriculectomy were analyzed for potential re-expression of MLC-2a at both the RNA and protein level.

Methods

Rat neonatal cardiomyocyte culture

Animals were handled according to the guidelines of the Animal Welfare Committee of the University of Maastricht. Cultured neonatal rat myocardial cells were prepared as previously described.[8] In brief, myocytes were isolated from the ventricles of 1-2 day old Wistar Kyoto rats by digestion with Collagenase II (Worthington, Lakewood NJ) and pancreatin (GibcoBRL, Paisley UK) at 37°C in a spinner flask. Purification of the cell suspensions was performed on a discontinuous Percoll gradient (Sigma Chemical Co, St Louis MO). The myocytes were plated in gelatinized 60 mm culture dishes at low density ($3.5x10^4$ cells/cm^2), in Dulbecco's modified Eagle Medium (DMEM) (Gibco BRL, cat no: 42430-025) and medium 199 (Gibco BRL, cat no: M199, 31153-026) 4:1 (v/v) supplemented with 10% horse serum and 5% fetal calf serum, (Sera-lab, Sussex England) and antibiotics (100 units/ml penicillin and 100 µg/ml streptomycin, GibcoBRL). Following overnight attachment in this medium the cells were washed twice with DMEM, and incubated with DMEM/medium199 (4:1; v/v) supplemented with glutamine and antibiotics. For myocyte stimulation studies the cells were incubated for 48 h in serum-free DMEM, containing glutamine and antibiotics in the presence of either 10^{-5} M phenylephrine (PE; Sigma), 10^{-8} M endothelin-1 (ET-1; Bachem, Bubendorf Switzerland), or β-estradiol 10^{-9} M (ICN pharmaceuticals, Costa Mesa CA). To evaluate the effects of serum, cells were maintained in 0%, 1%, or 10 % (v/v) new born calf serum (Sera-lab).

RNA isolation and Northern Blotting

RNA was isolated from rat neonatal cardiomyocytes using the TRIzol reagent (Gibco BRL). The quality of the RNA was routinely assessed by size fractionation on formaldehyde-agarose gels followed by ethidium bromide staining. Northern blot hybridizations were performed by minor modifications of a previously described method.[9] Briefly, the RNA (10 µg) was size fractionated by formaldehyde-agarose gel electrophoresis and transferred to nylon membranes (Hybond-N, Amersham Buckinghamshire UK) by capillary transfer. Filters were baked at 80°C for 30 minutes followed by a UV cross-linking of the RNA at 0.4 J/cm^2. Filters were prehybridized for 30 minutes at 65°C using a RapidHyb Buffer (Amersham). Subsequently, a 450 bp fragment of the MLC-2a cDNA, random labeled with α^{32}P dCTP (Dupont de Nemours NV, Brussels Belgium) was added to the hybridization solution and incubated for 2 h at 65°C. Stringent posthybridization wash conditions were used (0.1 x SSC, 0.1% SDS at 65°C).

Expression of the MLC-2a gene was detected by autoradiography and phosphor imaging. As a control, for loading and transfer efficiency, the filters were

routinely stained with methylene blue to visualize the 18 S and 28 S RNA bands. In addition hybridizations were performed using randomly labeled 18 S and ANF cDNA as probes. Quantification of Northern blot data was performed and is based on three independent myocyte isolations.

rt-PCR

First-Strand cDNA synthesis using 5 µg total RNA was performed using a poly-dT primer and the enzyme Reverse Transcriptase (Gibco BRL) in a volume of 20 µl. Following termination of the first-strand synthesis reaction, the samples were incubated with RNase H at 37 for 20 minutes. The same first-strand preparation was used for analyzing each of the gene products by PCR. PCR was performed with 2.5 µl of the first strand reaction, 1 unit of *Taq* polymerase (GibcoBRL) and 50 pmol of the appropriate primers in a reaction volume of 25 µl. The reactions were carried out in an automated thermal cycler. The amplification sequence consisted of an initial denaturation at 94°C for 2 min. Followed by 30 cycles of 94°C for 1 minute, 55- 65°C annealing temperature (depending on the melting temperature of the of the primers), and extension at 72°C for 1 minute. The final extension was done for 10 minutes at 72°C. The primers used for MLC-2a detection were sense 5'GCATGTTCGACCCCAG 3' and antisense 5' AGTTCAGGCACAGAG TTTATT 3' giving a PCR product of 240 bp. Quality of the First Strand cDNA was checked by PCR with murine GAPDH specific primers[10], giving yield to a PCR product of 300 bp. Primer sequence antisense 5'TTATTATAGGGTCTGGGATG 3' and sense 5'ACACTGAGGACCAGGTTGTC 3'.

Plasmid constructs and transfection by calcium phosphate precipitation

Using the convenient restriction sites in the 5' untranslated region of the MLC-2a gene deletion fragments were generated. (Pst I: 3.0 Kb, Pvu II: 1.8 Kb, Hinc II 0.8 Kb, Bam HI 0.2 Kb). Additional constructs containing 0.45 and 1.0 Kb of the MLC-2a regulatory DNA were generated by PCR and subcloning. All constructs were sequenced either at the ends to check for correct positioning or sequenced completely to exclude that PCR-induced mutations had been introduced. The fragments were cloned into the pXP1[11] or pGL2 (Promega, Madison WI) promoterless luciferase vector. Transient transfection of the myocytes was performed by the calcium phosphate precipitation method.[12] Each 60 mm culture dish of myocardial cells was transfected using 8 µg of a luciferase reporter vector and 2 µg of a cytomegalovirus-β-galactosidase expression vector. One day after transfection cells were washed twice with DMEM and were stimulated with PE, ET-1 or β-estradiol. The cells were harvested 48 h later. Control experiments were performed using 638 ANF-luciferase[13], 250 MLC-2v/luciferase[14] plasmids as the ANF and MLC-2v promoter have been shown to respond to hypertrophic stimuli.

To control for transfection efficiency the CMV-galactosidase (pON 249)[15] was used, while RSV-luciferase[16] served as a positive control. Data represent at least five experiments using three distinct cesiumchloride purified plasmid preparations.

Luciferase and β-galactosidase Assays

Cells were lysed in buffer containing (100 mM K_2HPO_4:KH_2PO_4 (92:8 vol.), pH 7.9, 0.5% Triton x-100, 1mM dithiothreitol) and assayed for β-galactosidase and luciferase activity. The final composition of the β-galactosidase reaction was 100 mM sodium phosphate, pH 7.3, 0.1 mM magnesium chloride, 2 mM magnesium sulphate, 40 mM β-mercaptoethanol and 4 mg/ml Ortho-Nitrophenyl-β-D-Galactopyranoside. β-Galactosidase activity in 20µl of cell extract was measured in triplicate at 420 nm in a Pharmacia-LKB Ultrospec-III spectrophotometer (Pharmacia Biotech AB, Uppsala Sweden). The final composition of the luciferase reaction medium was 67 µM D-Luciferin, 2 mM ATP, in 300 µl of KTME (100 mM Tricine, 10 mM $MgSO_4$, 2 mM EDTA, pH 7.8). Luciferase activity in 10 µl cell extract was measured in triplicate in a Perkin Elmer Luminesence Spectrometer LS 50 B (Perstorp Analytical Lumac, Landgraaf, the Netherlands). Data are presented as relative luciferase activity based on the luciferase/galactosidase (Luc/Gal) ratio.

Partial left ventriculectomy

After obtaining informed consent, tissue samples of two patients who underwent partial left ventriculectomy were analyzed for the expression of ANF and MLC-2a genes. Patients were selected for this procedure if they suffered from symptomatic heart failure (NYHA class III-IV), based on an idiopathic dilated cardiomyopathy (IDCM; left ventricular enddiastolic diameter by echocardiography > 70 mm). Patients with coronary artery disease (lesion > 50% in 2 vessels) were excluded. Additional exclusion criteria were previous or current adriamycin treatment, active myocarditis, muscular dystrophy, or a systemic illness with cardiac manifestations, including diabetes mellitus. In brief, cardioplegia is applied after aortic cross clamping. Standard extra corporal circulation is initiated. Posterolateral left ventricular tissue is resected between the papillary muscles, including the obtuse marginal branch. Resection is performed from apex to basis of the heart leaving the mitral valve intact. The heart is closed in two separate layers. The tissue is preserved on ice, dissected and frozen at -70°. For RNA and protein analysis we used tissue blocks of approximately 5 grams. For control experiments human atrial and murine ventricular tissue was obtained from one patient in sinus rhythm, undergoing coronary bypass surgery.

Western blotting

Flash frozen human biopsies were homogenized and protein was isolated in ice cold lysis buffer (0.5% non-idet P-40, 150 mM NaCl, 10 mM Tris-HCl pH 8.0, 5mM EDTA) in the presence of a cocktail of proteinase inhibitors (2 µg/ml aprotonin, 2 µg/ml pepstatin A, 2 µg/ml leupeptide, 10 µg/ml p-amidinophenulmethane-sulfonyl fluoride (Sigma) and 2 µg/ml soybean trypsin inhibitor (Gibco BRL). The amount of protein was estimated by the method of Bradford[17], using a Bio-Rad protein determination kit and bovine serum albumine (BSA, GibcoBRL) as a standard. For immunoblotting, proteins were separated on a 8% gel by SDS-PAGE, and transferred to nitrocellulose filter (Schleicher & Schuell, Keene NH). Filter was blocked for 1 hour at room temperature with phosphate buffered saline, containing 1% BSA and 0.2% non-idet P-40 (PBSAN). The first antibody used, was the polyclonal rabbit anti-mouse MLC-2a antibody (gift from Dr. Steve Kubalak, Charleston) diluted 1:500 in PBSAN and incubated overnight at 4°C. After three washes with PBSA, the secondary anti-rabbit immunoglobulin, peroxidase linked antibody (1:2500 dilution in PBSAN) was applied for 1 hour at room temperature. After three additional washes the membranes were subjected to procedures for Enhanced Chemiluminescence (Amersham), according to the supplier's instructions.

Data analysis

Luciferase/galactosidase (Luc/Gal) ratio was calculated to correct for differences in transfection efficiency within one experiment. To correct for variation in consecutive experiments the Luc/Gal ratios are expressed as relative to the Luc/Gal ratio of the 0.2 Kb MLC-2a promoter fragment, which was arbitrarily set at 1.0. The other results are represented as calculated activities relative to the 0.2 Kb fragment. For statistical analysis the non parametric Kruskal Wallis test was applied. P-values of <0.05 were considered to be statistically significant.

Results

MLC-2a expression in cultured ventricular myocytes

By Northern blotting and rt-PCR the expression of the MLC-2a gene in cultured ventricular cardiomyocytes was determined (figure 1A and B). The effect of various serum conditions on the level of expression was assessed. Even in the absence of serum (0%) expression of the MLC-2a gene was detected both by rt-PCR and Northern blotting. The highly sensitive rt-PCR approach did not reveal the presence of any MLC-2a message in total RNA from adult mouse

ventricular tissue, thereby validating the specificity of the PCR-conditions applied. The quality of the extracted RNA and rt-PCR reaction was comparable for the different conditions tested, as shown by the expression of the house keeping gene (glyceraldehyde-3-phosphate dehydrogenase: GAPDH). Stimulation of cardiomyocytes with PE and ET-1 showed a significant increase in the message for the MLC-2a gene as assessed by Northern blotting and this was supported by the rt-PCR results. This finding indicates activation of the endogenous promoter of the MLC-2a gene in ventricular cardiomyocytes. The strength of the endogenous ANF signal shows a similar pattern of induction at 1 and 10% serum addition and after PE, ET-1 and β-estradiol treatment. There was no marked change in the level of expression of ribosomal RNA (18S) detected by Northern blotting. Quantification of the Northern blot signal in three consecutive stimulation experiments revealed a five fold induction in MLC-2a expression levels upon stimulation with PE and ET-1. A similar increase (approximately five fold) was observed on endogenous MLC-2a expression in ventricular myocytes 48 hours after treatment with β-estradiol (figure 1C). Marked differences in myocyte appearance were observed under different experimental conditions. Cardio-myocytes maintained in the absence of serum or treated with β-estradiol remain small and cornered. Whereas in myocytes kept in medium with 10% serum no marked changes were observed going from plating medium to the experimental conditions. ET-1 and PE caused a marked increase in cardiomyocyte size. These observations were made by two independent investigators, in three different myocyte preparations.

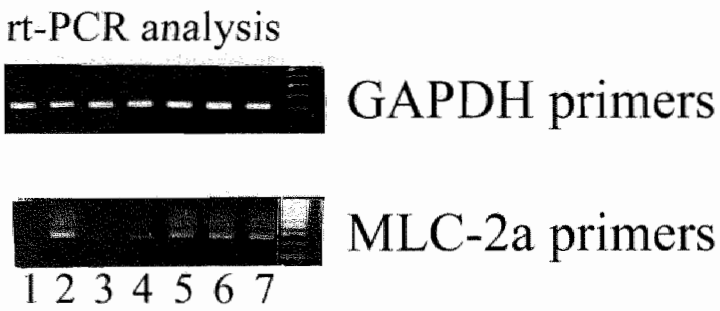


Figure 1A. rt-PCR for MLC-2a and GAPDH. No PCR product was generated with adult mouse ventricular RNA (lane 1). Total RNA from murine atrial tissue was used as a positive control (lane 2). Expected 250 bp PCR bands were found when using RNA from cultured ventricular myocytes, in the absence of serum (3), 10% serum (4), ET-1 (5), PE (6) and β-estradiol (7). To control the size of the PCR product a 100 bp DNA ladder was loaded in lane 8. The rt-PCR was applied here as a qualitative assay. The amount of PCR product does not necessarily correlate with the amount of mRNA template. The 300 bp murine GAPDH fragment was amplified to check the quality of the isolated total RNA.

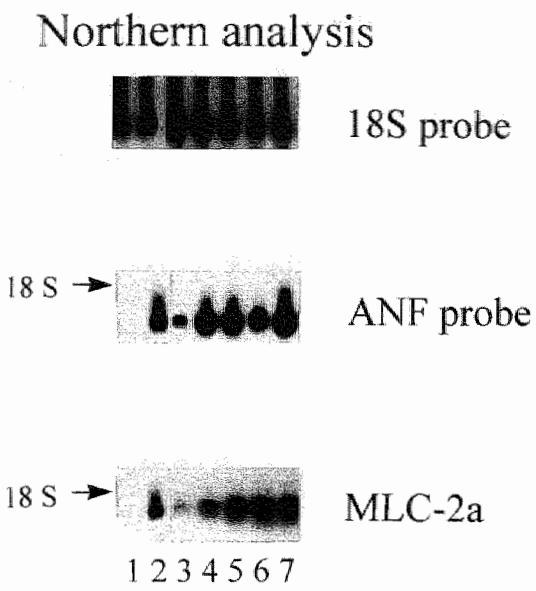


Figure 1B. Representative Northern Blot showing the effect of serum (0% lane 3 and 10% lane 4), ET-1 (5), PE (6), and β-estradiol (7) treatment on ventricular cardiomyocytes. In addition the level of ANF expression is shown. Hybridization signal with 18S probe shows that each lane contained about equal amounts of RNA.

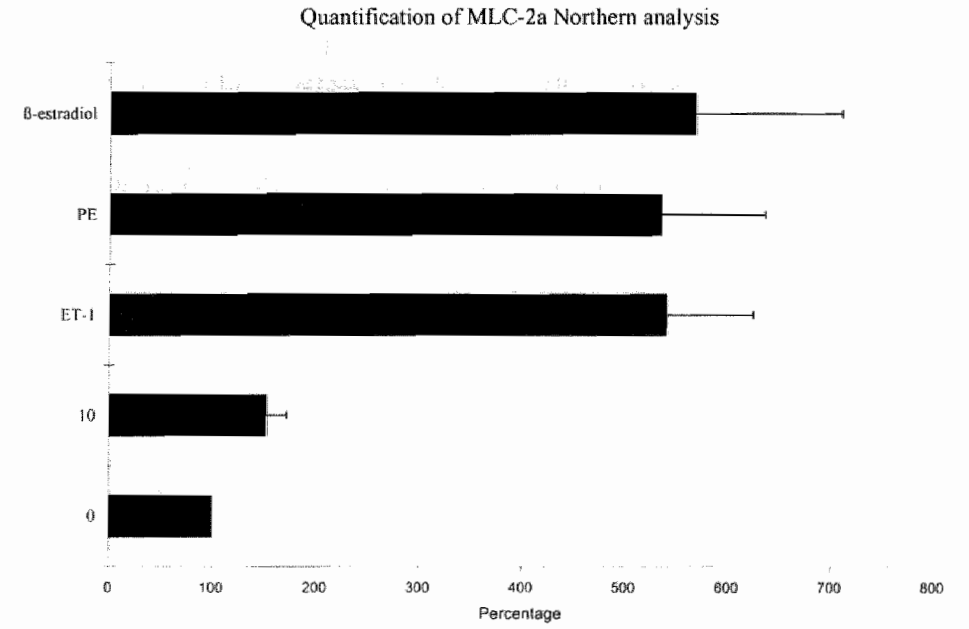


Figure 1C. Quantitative representation of Northern blot data. The MLC-2a signal was normalized using the 18S probe. Northern hybridization following β-estradiol, PE, ET-1 treatment and serum 0% and 10% addition was performed and measured in three independent cardiomyocyte cultures. The MLC-2a/18s signal in the absence of serum was set at 100%.

Transient transfection assays

Deletion fragments of the MLC-2a promoter driving luciferase were transfected in rat neonatal ventricular cardiomyocytes. The effect of serum addition (1 or 10%) was determined. The relative luciferase activity of the 0.2 Kb fragment was compared to the expression level of both the 638 ANF promoter-luciferase construct and the 250 MLC-2v promoter-luciferase construct. There was a trend (not significant) towards higher expression levels at higher serum conditions in all three constructs (figure 2A).

No significant difference was observed in relative luciferase activity normalized to the 0.2 Kb fragment, comparing 0% serum to 1 and 10%, for the MLC-2a promoter deletion constructs > 0.9 Kb tested. However the 0.45 Kb fragment showed a significantly higher expression level in the absence of serum and at 1% compared to 10% serum (figure 2B).
Cardiomyocytes were stimulated with β-estradiol, PE or ET-1 in the absence of serum.

PE stimulation resulted in approximately threefold increase in relative luciferase activity in the 1.8 and 3.0 Kb constructs (figure 3). The induction by ET-1 was more or less the same, ranging from 2 to 2.5-fold for the 1.8 and 3.0 Kb MLC-2a promoter fragments, respectively. In contrast no positive effect was observed after PE or ET-1 treatment of myocytes transfected with constructs < 1.8 Kb. β-Estradiol treatment of the cells resulted in a marked induction of relative luciferase activity, when the cells were transfected with promoter fragments > 1.8 Kb. The effect was most pronounced in the 1.8 Kb fragment (10x), compared with the 3.0 Kb fragment (7x). The stimulatory effect of β-estradiol was not observed with promoter fragments of 1.0 Kb or less (figure 3).

Analysis of human tissue samples

Tissue samples were obtained from two patients who underwent partial left ventriculectomy and both RNA and protein were isolated for Northern and Western blot analysis. The mouse cDNA was used as a probe to detect the expression of the human MLC-2a gene, using human atrial tissue as a positive control. In addition the blots were probed with a highly conserved fragment of the mouse ANF cDNA (figure 4). The Northern blot shows the high expression level of ANF in atrial and ventricular tissue of the two patients. A similar, albeit less pronounced expression pattern is detected for the MLC-2a message, which was detected in human atrium and patient left ventricular samples, but not in mouse ventricular tissue.

A polyclonal rabbit anti-mouse MLC-2a antibody is available and has been used successfully in rodent to distinguish atrial from ventricular cardiomyo-cytes.[18] In our study the antibody was used to detect (potential) translation of the

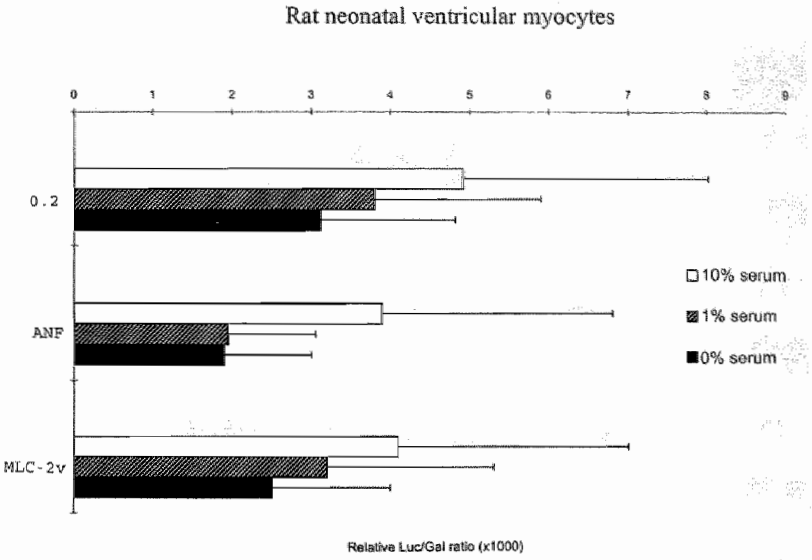


Figure 2A. Rat neonatal ventricular myocytes were transfected with the 0.2 Kb MLC-2a/luciferase, 0.25 Kb MLC-2v/luciferase, 0.638 Kb ANF/luciferase. The effect of no serum and the addition of either 1% or 10% (v/v) is shown. Data are shown as relative luciferase values (mean, ± sd).

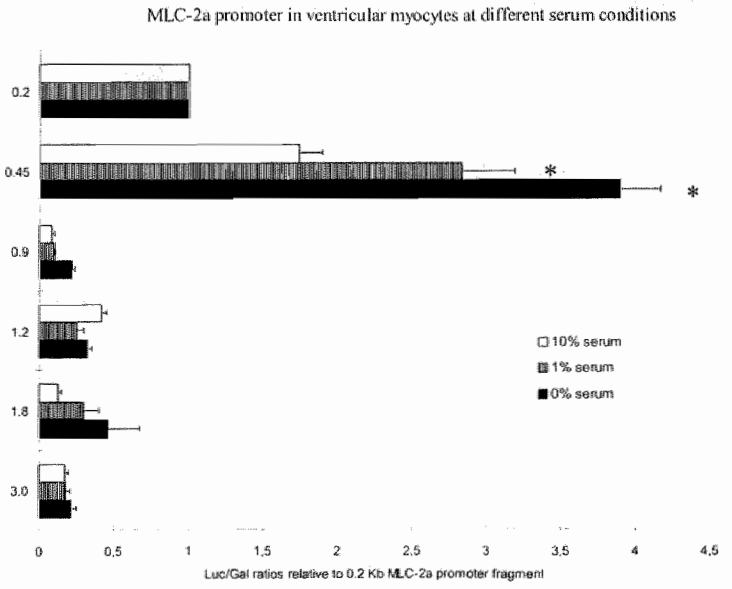


Figure 2B. Transfection results in rat neonatal cardiomyocytes with deletion fragments of the MLC-2a promoter varying in size from 0.2 Kb to 3.0 Kb, cultured after transfection in the absence of serum or at 1% or 10% serum (v/v). Data are shown is Luc/Gal ratios relative to the value measured for the matching 0.2 Kb fragment (set as 1.0). There is a significant increase in activity of the 0.45 Kb fragment at 1% and no serum conditions ($p < 0.05$ Kruskal Wallis).

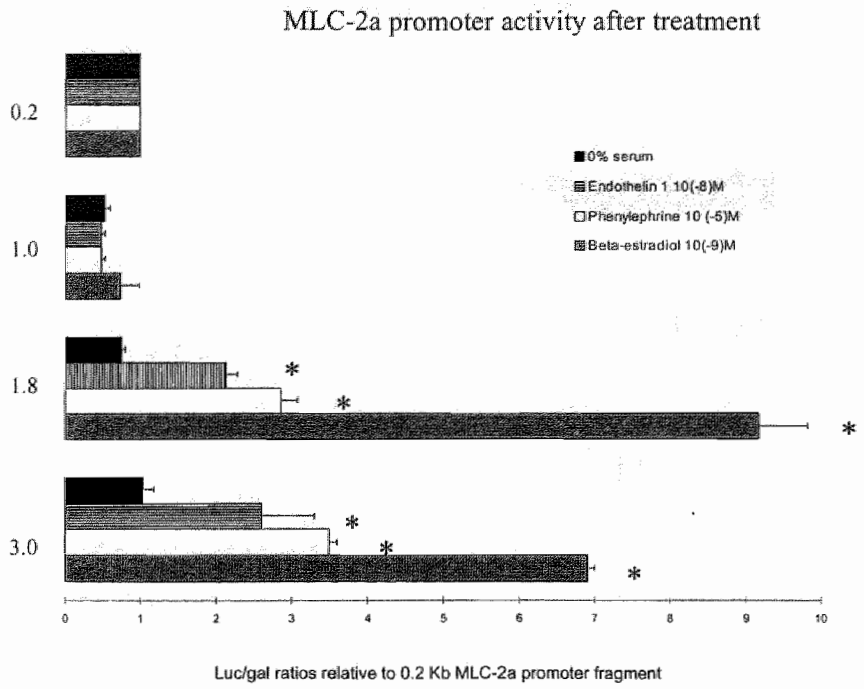


Figure 3. Results of transient transfection experiments in cultured rat neonatal ventricular cardiomyocytes. Different deletion fragments of the MLC-2a promoter were used varying in size form 0.2 to 3.0 Kb. Following transfection cells were incubated with medium in the absence of serum (0%), endothelin-1 (10^{-8}M), phenylephrine (10^{-5} M) or β-estradiol (10^{-9}M). There was a significant increase in the level of promoter activity for the 1.8 and 3.0 promoter fragments ($p < 0.05$ Kruskal Wallis). The relative luciferase activity for the 0.2 Kb fragment were (mean ± sd): no treatment, 2209 (±1102); endothelin-1, 2110 (±738); phenylephrine, 2359 (±978); β-estradiol, 1903 (±473).

MLC-2a mRNA into protein in human ventricular tissue. Western blotting showed that the murine antibody recognizes selectively the protein in atrial tissue, but not the ventricular isoform in murine ventricular tissue. Furthermore in left ventricular tissue of patients with IDCM re-expression of MLC-2a protein was demonstrated.

Discussion

The MLC-2a gene is re-expressed in cultured rat neonatal ventricular myocytes. The MLC-2a mRNA is detectable when the cells are maintained in the absence of serum but levels increase when either 10% serum, ET-1, PE or β-estradiol is added to the medium. Transient transfection studies with deletion fragments of the MLC-2a 5' flanking region reveal that the region spanning from -1.0 to -1.8

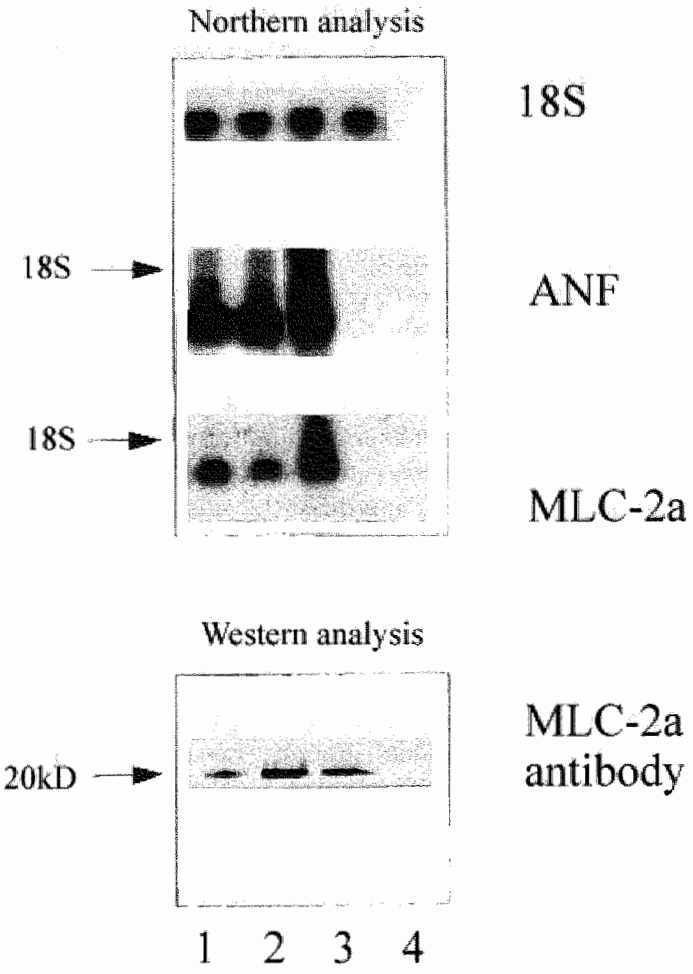


Figure 4. Northern blotting: Total RNA from ventricular tissue of patients undergoing partial left ventriculectomy (lane 1 and 2), human right atrial appendix tissue (lane 3), adult murine left ventricular free wall. Mouse ventricular tissue was used as negative control as there was no control human ventricular tissue sample available. Shown in panel A, top 18S hybridization, to show equal loading of the filter, Atrial Natriuretic Factor and MLC-2a. No signal was found for ANF and MLC-2a in mouse ventricular tissue and as expected ANF and MLC-2a are detected in human atrium. There is a marked level of expression in both patients in cardiac failure for ANF and MLC-2a. Western blotting: Protein extracted from left ventricular tissue samples (1,2) human atrium (3), adult murine ventricular tissue (4). The primary antibody was polyclonal rabbit anti-mouse MLC-2a (1:500), the secondary antibody is peroxidase labeled anti-rabbit immunoglobin (1:2500).

Kb, is responsible for upregulating reporter gene (luciferase) activity upon treatment of myocytes with ET-1, PE and β-estradiol. In human left ventricular cardiac tissue of patients with IDCM we found the expression of the MLC-2a gene both at the mRNA and protein level.

Re-expression of the MLC-2a gene in cultured ventricular myocytes

The endogenous MLC-2a gene is expressed in rat neonatal ventricular cardiomyocytes in defined serum-free medium, whereas in neonatal cardiac tissue *in vivo* expression of MLC-2a is exclusively found in atrial myocytes. The re-expression of the MLC-2a gene shown in this study suggests that the procedure of isolating and culturing ventricular cardiomyocytes in itself provides the stimulus for re-expression.
Several factors could be important for the induction of MLC-2a expression under these conditions. Myocyte isolation is traumatic to the cell as the myocytes are separated from neighboring cells and the extracellular matrix becomes degraded. The cardiomyocytes are plated at a relative low density, so that there is basically no restoration of cell/cell contact possible. In addition the cells will have to attach to the culture dishes to survive the procedure. In theory each of these factors could be responsible for the induction of MLC-2a expression. It should be noted that the re-expression of a fetal gene like MLC-2a is not necessarily a reflection of the activation of the complete "hypertrophy program". Although cardiomyocytes maintained in serum-free medium express "markers" of hypertrophy, the cells displayed a small and cornered morphological appearance. Similar observations have been made in cardiomyocytes treated with β-estradiol. In contrast, the ET-1, PE and 10% serum treated cells appear much larger. The neonatal rat ventricular myocytes maintained in serum free medium with or without β-estradiol treatment show MLC-2a expression without morphologic signs of myocyte hypertrophy.

The re-expression of fetal genes in cultured cardiomyocytes in the absence of serum, was shown by Wollert et al.[4] for the ANF and α skeletal actin gene. Total RNA from neonatal rat cardiomyocytes was analyzed after a 48 hour incubation period in serum free medium. No morphologic signs of hypertrophy were reported for these cells. In fact the measurements in the 'serum-free' cardiomyocytes provide the baseline measurement for hypertrophy inducing experiments.[4,19] In our experiments treatment of cardiomyocytes with β-estradiol showed a high level of MLC-2a expression without overt physical hypertrophy. A similar effect of estrogen treatment on gene expression in cardiomyocytes has not been reported previously.

Re-expression of the MLC-2a gene after ET-1 and PE treatment

Treatment of cultured neonatal rat cardiomyocytes with PE and ET-1 showed increased MLC-2a expression concordant with morphological signs of hypertrophy. ET-1 and PE cardiomyocyte treatment results in expression of immediate early genes, the induction of contractile protein synthesis,[20] and the reactivation of a fetal gene program.[13,21-26] Different receptor-signaling pathways in ventricular cardiomyocytes are involved in sensing hypertrophic stimuli.[2] Endothelin receptor A is the predominant receptor isoform found in cardiomyocyte membranes[24], and PE activates the α-adrenergic receptors. Endothelin receptor A and α-adrenergic receptors have been shown to interact with G-proteins. For both growth factors signaling involves activation of the Ras/Raf pathway, while several other routes of intracellular signaling can be used.[27] In general the signal transduction pathways connecting the receptor for a hypertrophic stimulus to transcriptional regulation are diverse and highly complexed through cross-talk between the pathways involved (review[28]).

Although there are many similarities in the hypertrophic response induced by ET-1 and PE, several reports indicate that different fetal gene expression programs exist and morphologic signs of hypertrophy can vary. The expression of the endogenous ANF gene in relation to hypertrophy was studied by Glembotski and coworkers.[26] They showed that during prolonged stimulation with PE, the expression of ANF was maintained for more than 72 hrs at a constant level, whereas ET-1 gave a transient rise in ANF levels, while both treatment groups exhibit morphologic signs of cardiomyocyte hypertrophy.

Potentially by altering either protein level or protein activity the significance of key players in signal transduction pathways can be studied in cultured cardiomyocytes. Performing such signal transduction interventions studies, the dissociation between fetal gene expression and morphologic signs of hypertrophy was underscored by the findings of Fuller et al.[29] By over-expressing a phosphatase (Mitogen-activated protein kinase phosphatase 1:MKP1) they blocked the re-expression of fetal genes on hypertrophic stimuli without a significant reduction of morphologic hypertrophy. MKP1 can inactivate mitogen activated protein kinase, a key protein in the Ras/Raf signaling pathway.[27]

Cardiotrophin-1 (CT-1) treatment of cardiomyocytes induces a different form of hypertrophy. The CT-1 protein belongs to a family of cytokines requiring the receptor protein gp-130 to mediate its effect. CT-1 induces expression of ANF but not skeletal α-actin. In contrast stimulation by PE showed elevated levels of both ANF and skeletal α-actin. In these studies the potential re-expression of the MLC-2a gene was not studied. PE treatment gave widening and lengthening of cardiomyocytes, whereas the CT-1 induced hypertrophy was characterized only by lengthening of the cells.

The situation *in vivo* is more complex. Ventricular myocyte selective overexpression of the constitutively active Ras gene (dominant positive) in transgenic mice resulted in a hypertrophic phenotype with intracavitary flow obstruction. Molecular analysis of this mouse model of hypertrophic obstructive cardiomyopathy showed a selective upregulation of atrial and brain natriuretic factor, without increased levels of contractile protein expression (including MLC-2a).[30] These findings are in line with the *in vitro* results with CT-1 stimulation. In contrast re-expression of MLC-2a in the ventricles of aortic banded wild type mice was demonstrated by RNA analysis. The results obtained in the transgenic mice and aortic banded mice indicate that it is not the Ras/Raf signaling cascade *per se*, which is responsible for ventricular re-expression of the endogenous MLC-2a in mice *in vivo*.

Exogenous MLC-2a promoter activation in ventricular myocytes:

To determine whether the activation of the endogenous MLC-2a gene to serum is mediated by the cloned 5' flanking region of the gene, transient transfections were performed with MLC-2a promoter deletion fragments linked to the luciferase reporter gene. The level of expression obtained with the 638 ANF promoter and 250 MLC-2v promoter is comparable with the 0.2 Kb MLC-2a promoter (figure 2A). The behavior of the MLC-2v and ANF promoter was as anticipated, showing an increase upon serum addition.[23] There is a discrepancy in the response of the 0.2 Kb fragment compared to the 0.45 Kb fragment, when the cardiomyocytes are incubated in the presence or absence of serum. The 0.2 Kb activity construct shows a trend to increased MLC-2a, while the 0.45 Kb is progressively repressed with increasing serum concentration (figure 2B).

The 0.45 Kb promoter fragment has several E-boxes, one GATA and one retinoic acid responsive element in addition to the conserved *cis* elements of the 0.2 Kb fragment.

Undefined factors in serum can reduce the expression level of the 0.45 Kb fragment. Still at all serum concentrations, higher expression levels were obtained with the 0.45 Kb construct. The relatively strong promoter activity of this fragment can be important for future transgenic experiments.[31]

The 0.9, 1.0 and 1.2 Kb constructs are less active compared to the 0.2 Kb fragment. This repressive effect is not dependent on serum conditions or neurohormonal stimulation and suggestive for the presence of a transcriptional silencer in the region between 0.45 Kb to 0.9 Kb (chapter 7).

The MLC-2a promoter fragments of 1.8 and 3.0 Kb are strongly activated by ET-1 and PE treatment, as assessed by the increased luciferase activity. The transfection data indicate the presence of cis elements responsible for ET-1 and PE induction between position -1.0 Kb to -1.8 Kb. Sequence comparison revealed, amongst

others, two AT-rich elements within this promoter fragment that may act as potential myocyte enhancer factor (MEF2) binding sites (chapter 7).

The stimulating effect of PE treatment on promoter activity of ANF and ventricular MLC-2 in transient transfection studies, can be mimicked by application of constitutively active Ras. The pathway is blocked in the presence of a constitutively active mitogen activated protein kinase (MAPKK). The *cis* element of the MLC-2v promoter involved in the MAPKK signal transduction was mapped to the HF1B site, also an AT- rich element with which MEF2 can interact.[32] The transgenic experiment with the constitutively active RAS gene, shows that the behavior of the endogenous complete MLC-2v promoter is different.[30] The role of the AT-rich elements in the MLC-2a promoter will need to be addressed in future studies.

Estrogen activation of the MLC-2a promoter

β-Estradiol treatment of rat neonatal ventricular myocytes, transfected with the long promoter fragments (1.8 and 3.0 Kb) resulted in a more than 7-fold increase in relative luciferase activity. The DNA site required for estrogen induction has been mapped to the fragment from -1.0 to -1.8 Kb of the MLC-2a promoter. The effect of estrogen treatment on the endogenous MLC-2a gene was evaluated by Northern blotting, and revealed an approximately 5-fold increase. These findings point to an important positive regulating effect of estrogen on the transcription level of the MLC-2a promoter.

The presence of functional estrogen receptors in ventricular myocytes has been shown by Grohe et al.[33] Estrogen receptors belong to the superfamily of nuclear receptors or more specifically to the group of the steroid hormone receptors (Class 1).[34,35] The activated receptors translocate to the nucleus and homodimerize after ligand (estrogen) binding. The ligand-receptor complex binds to a palindromic sequence AGGTCA with a spacing of three nucleotides. After ligand-receptor interaction with the *cis* element of the target gene an interaction with the transcription initiation complex has been shown to occur. Through this interaction the level of transcription can be upregulated.[35] Sequence comparison of the MLC-2a promoter fragment from -1.0 to -1.8 Kb revealed a conserved hormone responsive element, with an imperfect palindrome spaced by three bases at position -1622 (TCCAGAggcTCTGGT). The physiological significance of this finding is unclear. Sofar the effect of estrogens has been studied epidemiologically[36], in oncology and vascular wall physiology. It has been shown that estrogen treatment has beneficial effects on vascular tone, vascular resistance and cardiac output in monkeys.[37,38] Whether estrogen treatment induces the expression of MLC-2a in ventricular tissue *in vivo* is presently unknown. The behavior of the transfected MLC-2a promoter is very interesting

and provides a unique tool to unravel the signaling pathway involved in estrogen-receptor interaction and DNA binding in cardiomyocytes.

MLC-2a expression in dilated cardiomyopathy

Using murine cDNA and antibodies, we documented re-expression of the MLC-2a gene in ventricular tissue of selected patients with cardiac failure due to IDCM. Simultaneously, we showed re-expression of ANF at the mRNA level.

There are few reports on the behavior of the cardiac myosin light chains in pathologic conditions. One study reports a marked decrease in ventricular MLC-2 protein level, due to the activation of a protease selectively degrading the ventricular MLC-2 protein in IDCM.[39] Other investigators found no difference at the mRNA level of the ventricular MLC-2 gene comparing tissue from normal with IDCM hearts. The atrial MLC-1 gene is re-expressed in ventricular tissue of IDCM and post-myocardial infarction.[39,40] None of these studies looked for re-expression of the MLC-2a gene in human ventricular tissue.

It is uncertain whether the re-expression of the MLC-2a in the human ventricle is beneficial for the ventricular function. Interesting data from isoform substitutions in transgenic mice were recently published. Palermo and co-workers were able to replace the MLC-2a protein in the atrial chambers completely with MLC-2v.[41] They performed hemodynamic measurements on isolated *ex vivo* working hearts from these transgenics and showed a diminished cardiac output.[42] They concluded that the substitution of atrial isoform is deleterious for atrial function and causes a reduction in cardiac output. The reverse experiment to replace MLC-2v by MLC-2a in the ventricles has not been performed yet.

The striking finding in the transgenic MLC substitution studies is the tight control in sarcomeric protein stoichiometry. For instance in transgenic mice overexpressing the MLC-2v gene, the MLC-2v protein level is identical to non-transgenic mice, despite marked increased of MLC-2v mRNA levels. In a similar experiment Gulick et al, introduced the fast muscle specific MLC-2f as a transgene.[43] The replacement of the MLC-2a protein in atrial tissue was efficient resulting in nearly complete (>90%) replacement by MLC-2f. In contrast the ventricular tissue was resistent to isoform replacement resulting in a maximal MLC-2v protein replacement of 55%. Functional analysis of the hearts of these transgenic mice revealed a modest though significant decrease in ventricular contractility and impaired relaxation. These results suggest differences in affinity of the various MLC-2 isoforms for the contractile apparatus, where the MLC-2a protein has a low and the MLC-2v displays a strong affinity.

By Western blot analysis the re-expression of MLC-2a in left ventricular tissue of patients with dilated cardiomyopathy was shown. Whether the human MLC-2a protein is incorporated in the sarcomeric structures will have to be determined by immunohistochemistry. Furthermore it will be interesting to study contrac-

tility in left ventricular tissue samples or isolated ventricular myocytes. At present the functional implications of our findings are still unknown.

Conclusions

The endogenous MLC-2a gene is re-expressed in neonatal rat ventricular cardiomyocytes and its expression is further enhanced by hypertrophic stimuli, like PE and ET-1. A similar response was found in transient transfection assays using the MLC-2a promoter. Two AT rich *cis* elements reside in the activated promoter fragment and may be involved in ET-1 and PE-mediated transcription activation. β-Estradiol showed a strong induction of transfected MLC-2a promoter fragment activity. Furthermore a five fold upregulation of the endogenous MLC-2a gene was found. A palindromic HRE has been identified on the estrogen responsive MLC-2a promoter fragment. In addition the re-expression of the MLC-2a gene in patients with dilated cardiomyopathy was demonstrated. Whether the MLC-2a re-expression is due to hypertrophy *per se* or related to the transition to cardiac failure remains to be established.

References

1. Grossman W: Cardiac hypertrophy: adaptation or pathologic process. *Am J Med* 1980;576-583
2. Lembo G, Hunter JJ, Chien KR: Signaling pathways for cardiac growth and hypertrophy. Recent advances and prospects for growth factor therapy. *Ann N Y Acad Sci* 1995;752:115-127
3. van Bilsen M, Chien KR: Growth and hypertrophy of the heart: towards an understanding of cardiac specific and inducible gene expression. *Cardiovasc Res* 1993;27:1140-1149
4. Wollert KC, Taga T, Saito M, Narazaki M, Kishimoto T, Glembotski CC, Vernallis AB, Heath JK, Pennica D, Wood WI, Chien KR: Cardiotrophin-1 activates a distinct form of cardiac muscle cell hypertrophy. Assembly of sarcomeric units in series VIA gp130/leukemia inhibitory factor receptor-dependent pathways. *J Biol Chem* 1996;271:9535-9545
5. Neyses L, Vetter H: Molecular biology of oncogenes and cardiovascular hypertrophy. *J Hypertens* 1992;10:1447-1452
6. Kubalak SW, Miller-Hance W, O'Brien T, Dyson E, Chien K: Chamber-specification of atrial myosin light chain-2 expression precedes septation during mouse cardiogenesis. *J Biol Chem* 1994;269:16961-16970

7. Rockman H, Knowlton KU, Ross J, Jr., Chien KR: In vivo murine cardiac hypertrophy: a novel model to identify genetic signaling mechanisms that activate an adaptive physiological response. *Circulation* 1993;VII 14-21
8. De Vries JE, Vork MM, Roemen THM, de Jong YF, Cleutjens JPM, van der Vusse GJ, van Bilsen M: Saturated, but not mono-unsaturated fatty acids induce apoptotic cell death in neonatal rat ventricular myocytes. *J Lipid Res* 1997;38:1384-1394
9. Iwaki K, Sukhatme VP, Shubeita HE, Chien KR: Alpha-and beta-adrenergic stimulation induces distinct patterns of immediate early gene expression in neonatal rat myocardial cells. fos/jun expression is associated with sarcomere assembly; Egr-1 induction is primarily an alpha 1-mediated response. *J Biol Chem* 1990;265:13809-13817
10. Sabath DE, Broome HE, Prystowsky MB: Glyceraldehyde-3-phosphate dehydrogenase mRNA is a major interleukin 2 induced trnascript in a cloned T-helper lymphocyte. *Gene* 1990;91:185-191
11. Nordeen SK: Luciferase reporter gene vectors for analysis of promoter and enhancers. *Biotech* 1988;6:454-457
12. Chen C, Okayama H: High-efficiency transformation of mammalian cells by plasmid DNA. *Mol Cell Biol* 1987;7:2745-2752
13. Knowlton KU, Baracchini E, Ross RS, Harris AN, Henderson SA, Evans SM, Glembotski CC, Chien KR: Co-regulation of the atrial natriuretic factor and cardiac myosin light chain-2 genes during alpha-adrenergic stimulation of neonatal rat ventricular cells. Identification of cis sequences within an embryonic and a constitutive contractile protein gene which mediate inducible expression. *J Biol Chem* 1991;266:7759-7768
14. Zhu H, Garcia AV, Ross RS, Evans SM, Chien KR: A conserved 28-base-pair element (HF-1) in the rat cardiac myosin light-chain-2 gene confers cardiac-specific and alpha-adrenergic-inducible expression in cultured neonatal rat myocardial cells. *Mol Cell Biol* 1991;11:2273-2281
15. Cherrington JM, Mocarski ES: Human cytomegalovirus transactivates the a promoter-enhancer via an 18-base pair repeat element. *J Virol* 1989;63:1435-1440
16. de Wet JR, Wood KV, DeLuca M, Helinski D, Subramani S: Firefly luciferase gene: structure and expression in mammalian cells. *Mol Cell Biol* 1986;7:725-737
17. Bradford MM: A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Annal Biochem* 1976;72:248-254
18. Kubalak S, Doevendans PA, Rockman H, Hunter JJ, Tanaka N, Ross Jj, Chien KR: Molecular analysis of cardiac muscle diseases based on mouse genetics, in Adolph KW (ed): *Human molecular genetics*. Orlando, Academic Press, 1996, pp 470-487
19. Zhou MD, Sucov HM, Evans RM, Chien KR: Retinoid-dependent pathways suppress myocardial cell hypertrophy. *Proc Natl Acad Sci U S A* 1995;92:7391-7395

20. Shubeita HE, McDonough PM, Harris AN, Knowlton KU, Glembotski CC, Brown JH, Chien KR: Endothelin induction of inositol phospholipid hydrolysis, sarcomere assembly, and cardiac gene expression in ventricular myocytes. A paracrine mechanism for myocardial cell hypertrophy. *J Biol Chem* 1990;265:20555-20562
21. Sei CA, Glembotski CC: Calcium dependence of phenylephrine-, endothelin-, and potassium chloride-stimulated atrial natriuretic factor expression form long term primary neonatal rat atrial cardiocytes. *J Biol Chem* 1990;265:7166-7172
22. Soonpaa MH, Field LJ: Assessment of cardiomyocyte DNA synthesis in normal and injured adult mouse hearts. *Am J Physiol* 1997;272:H220-6
23. Shubeita HE, Martinson EA, van Bilsen M, Chien KR, Brown JH: Transcriptional activation of the cardiac myosin light chain 2 and atrial natriuretic factor genes by protein kinase C in neonatal rat ventricular myocytes. *Proc Natl Acad Sci U S A* 1992;89:1305-1309
24. Hilal Dandan R, Merck DT, Lujan JP, Brunton LL: Coupling of the type A endothelin receptor to multiple responses in adult rat cardiac myocytes. *Mol Pharmacol* 1994;45:1183-1190
25. Brown LA, Nunez DJ, Brookes CI, Wilkins MR: Selective increase in endothelin-1 and endothelin A receptor subtype in the hypertrophied myocardium of the aorto-venacaval fistula rat. *Cardiovasc Res* 1995;29:768-774
26. McDonough PM, Brown JH, Glembotski CC: Phenylephrine and endothelin differentially stimulate cardiac PI hydrolysis and ANF expression. *Am J Physiol* 1993;264:H625-30
27. Thorburn A, Thorburn J, Chen SY, Powers S, Shubeita HE, Feramisco JR, Chien KR: HRas-dependent pathways can activate morphological and genetic markers of cardiac muscle cell hypertrophy. *J Biol Chem* 1993;268:2244-2249
28. van Bilsen M: Signal transduction revisited. *Cardiovasc Res* 1997;In press:
29. Fuller SJ, Davies EL, Gillespie Brown J, Sun H, Tonks NK: Mitogen-activated protein kinase phosphatase 1 inhibits the stimulation of gene expression by hypertrophic agonists in cardiac myocytes. *Biochem J* 1997;323:313-319
30. Gotschall K, Hunter JJ, Tanaka N, Dalton N, Becker KD, Ross J, Jr., Chien KR: Ras-dependent pathways induce obstrucitive hypertrophy in echo-selected transgenic mice. *Proc Natl Sci USA* 1997;94:4710-4715
31. Doevendans PA, Hunter JJ, Lembo G, Wollert KC, Chien KR: Strategies for studying cardiovascular diseases in transgenic mice and gene-targeted mice, in Monastersky GM, Robl JM (eds): *Strategies in transgenic animal science.* Washington, American Society For Microbiology, 1995, pp 107-144
32. Thorburn J, Carlson M, Mansour SJ, Chien KR, Ahn NG, Thorburn A: Inhibition of a signaling pathway in cardiac muscle cells by active mitogen-activated protein kinase kinase. *Mol Biol Cell* 1995;6:1479-1490
33. Grohe C, Kahlert K, Briesemeister G, Stimpel M, Vetter H, Neyses L: Myocardial and myogenic cells contain functional estrogen receptors. *Circulation* 1994;90:2898 (abstract)

34. Mangelsdorf DJ, Evans RM: The RXR heterodimers and orphan receptors. *Cell* 1995;83:841-850
35. Beato M, Herrlich P, Schutz G: Steroid hormone receptors: many actors in search of a plot. *Cell* 1995;83:851-857
36. Grodstein F, Stampfer MJ, Colditz GA, Willet WC, Manson JE, Joffe M, Rosner B, Fuchs C, Hankinson SE, Hunter DJ, Hennekens CH, Speizer FE: Postmenopausal hormone therapy and mortality. *N Engl J Med* 1997;336:1769-1775
37. Pelzer T, Shamim A, Neyses L: Estrogen effects in the heart. *Mol Cell Biochem* 1996;160-161:307-313
38. Williams JK, Kim YD, Adams MR, Chen MF, Myers AK, Ramwell PW: Effects of estrogen on cardiovascular responses of premenopausal monkeys. *J Pharmacol Exp Ther* 1994;271:671-676
39. Trahair T, Yeoh T, Cartmill T, Keogh A, Spratt P, Chang V, dos Remedios CG, Gunning P: Myosin light chain gene expression associated with disease states of the human heart. *J Mol Cell Cardiol* 1993;25:577-585
40. Fujimoto K, Yasue H, Hashida S, Nakao K, Ishikawa E, Miyamoto E: Augmented expression of atrial myosin light chain 1 in ventricular aneurysms of human: enzyme immunoassay for atrial myosin light chain 1. *Biochem Biophys Res Commun* 1995;207:75-79
41. Palermo J, Gulick J, Colbert M, Fewell J, Robbins J: Transgenic remodeling of the contractile apparatus in the mammalian heart. *Circ Res* 1996;78:504-509
42. Palermo J, Gulick J, Ng W, Grupp IL, Grupp G, Robbins J: Remodeling the mammalian heart using transgenesis. *Cell Mol Biol Res* 1995;41:501-509
43. Gulick J, Hewett TE, Klevitsky R, Buck SH, Moss RL, Robbins J: Transgenic remodeling of the regulatory myosin light chains in the mammalian heart. *Circ Res* 1997;80:655-664

Chapter

Inhibition of the atrial myosin light chain 2 promoter in cultured cardiomyocytes by nuclear receptors

Pilar Ruiz-Lozano, Esther Merki, Pieter A Doevendans, Ming Dong Zhou, Kenneth R. Chien

Abstract

Background: Cardiac malformations occur in vitamin A deficiency and in the genetic model of retinoid X receptor alpha (RXRα) homozygous knock out mice. An aberrant persistent expression in ventricles of the atrial myosin light chain 2 (MLC-2a) is observed in retinoid deficiency, suggesting a direct or indirect effect of RXRα on transcriptional regulation of this gene.
Methods: Transfection studies in rat neonatal ventricular and atrial cardiomyocytes were performed, and the effect of pharmacological treatment with retinoic acid, and RXRα agonists was studied. In addition DNA-protein binding studies were done to verify the interaction of a conserved retinoic acid responsive element (DR-1) of the atrial myosin light chain 2 promoter with various members of the nuclear receptor superfamily.
Results: In transfection studies we determined that a MLC-2a/luciferase reporter gene is down-regulated after treatment with RXR-specific agonists in a dose-dependent manner. No significant response was observed after treatment with all-trans retinoic acid or with the RAR-specific agonist TTNBP. MLC-2a downregulation is blunted after mutation of DR-1 (mlcDR-1), that is dependent on its promoter context. In gel mobility shift assays, mlcDR-1 discriminates between RXR/RXR, RXR/RAR and RXR/PPAR heterodimer binding. Treatment with nuclear receptor selective agonists demonstrate that RXR and PPAR cooperate to regulate MLC-2a transcription.
Conclusions: These results provide evidence for a direct role of RXR in the regulation of chamber-specific gene expression, as RXR activation can downregulate MLC-2a promoter activity in cultured ventricular cardiomyocytes.

Introduction

Nuclear receptors are ligand-dependent transcription factors that regulate gene pathways involved in controlling growth, reproduction and homeostasis.[1-3] Ligand-dependent activation of nuclear receptors is also required for cell growth and differentiation [4,5] The ligand-activated nuclear receptors differentially regulate gene expression through their ability to recognize specific ligand responsive 'half-site' DNA elements[3,6] in the promoter sequence of target genes where the receptors can bind as heterodimers, homodimers or monomers.[7] Half-site sequences are typically arranged as direct repeats (DR), inverted or everted repeats and spaced by a variable nucleotide sequence responsible for the specificity of binding by different heterodimers.[8,9] Differential response element recognition distinguish four subclasses within the nuclear receptor super-family.[7,10,11] Class I are receptors for steroid hormones that bind as homodimers to inverted repeat (palindromic) response elements. Class II heterodimerize with RXR and bind to direct or symmetrical repeats. Receptors of class III bind to direct repeats as monomers and receptors of class IV bind to extended core sites as monomers.[12] Differences either in polarity or in sequence of the half site and the extension of the response element can have important effects on the activity of a determined receptor.[13,14]

Among heterodimeric nuclear receptors, retinoid X receptor (RXR) plays a central role as the obligate partner for heterodimerization. Heterodimeric partners for RXR include retinoic acid receptors (RARs), thyroid hormone receptors (TRs), vitamin D receptors (VDRs) and peroxisomal proliferator-activated receptors (PPARs). [12] The ability of RXR to activate or repress the transcription of target genes is thought to be differentially determined by its dimerization partners.[14-17] Knockouts for the nuclear receptors have led to significant advances in the understanding of their function.[18,19] In particular, RXRα null mutant (homozygous knock out) mice display malformations in the eye, defects in the ventricular compact zone and in cushion tissue mesenchyme.[18,20,21] RXRα mutant mice show persistent expression in the ventricle of the atrial marker myosin light chain 2a (MLC-2a)[22] and display features of an atrial-like phenotype, including thinning of the ventricular wall and decreased trabeculation.[22]

The failure of MLC-2a to be down-regulated in ventricles of the RXRα knockout mice suggests that retinoid signaling mediated by RXRα is required to allow progression of ventricular development from its early atrial-like phenotype to the adult ventricle.[22] The molecular mechanisms underlying atrial-like phenotype in the RXRα null mutant mouse may imply a change in the transcriptional program of ventricular cells. Here we report that the MLC-2a promoter contains functional retinoic acid response elements (RARE), whose activation down-regulate a MLC-2a/luciferase reporter gene in ventricular myocytes. By site directed mutagenesis, we have mapped retinoid-mediated repression to a DR-1

RARE, named mlcDR-1. Subsequently, we identified the binding activity of various nuclear receptors to mlcDR-1 and correlated this binding element specifically with functional effects of nuclear receptor-selective agonists. Our results suggest a direct role for RXR together with the peroxisome proliferator-activated receptors (PPARs) in the negative regulation of the atrial MLC-2 gene.

Methods

Plasmid constructs: Plasmid 1MLC-2a-Luc contained a 1030bp fragment of the MLC-2a promoter subcloned into the Xho I/Kpn I sites of a pGL2-luciferase vector (Promega, Madison WI). Deletion and mutational analysis were performed by PCR using the primers indicated below. The resulting PCR fragments were subcloned into Xho I and Kpn I sites of pGL2 vector. Oligos for 0.93MLC2a-Luc were 5'-CCGGGGTACCTAGTCAGAGACCTACTCATG-3' and 5'-CCCCC-TCGAGTGTGGCGGTCTGACTGAGTG-3'. mt1MLC-2a was generated by ligation of two fragments (1 and 2) using a Pst I site added to the primers. This site replaced the DR-1 element. Oligonucleotides to generate fragment 1 were: 5'-CCGGGGTACCGTGGGACTTTCCTAACAGG-3' and 5'-CCGGCTGCAGTCCTC-ACCCCAATGTGGCTAT-3'. To generate fragment 2: 5'-CCGGCTGCAGTAGT-CAGAGACCTACTCATG-3' and 5'-CCCCCTCGAGTGT-GG CGGTCTGACTGAG-TG-3'. Construct DR-1pt109 was generated upon subcloning one copy of a synthetic oligonucleotide containing -946 to -927 of the MLC-2a promoter into Hind III and Kpn I sites of the vector pT109 which contains the thymidine kinase basal promoter driving luciferase.[23] Expression vectors cmvRAR, cmvRXR and cmvRAR403 are described elsewhere.[24] In addition the luciferase reporter constructs, CRBP-II/luciferase and βRE-I/luciferase, containing respectively DR-1 and DR-5 RARE, in a truncated thymidine kinase promoter (pT109) were used.[25,26]

Tissue culture and transfections: Neonatal rat (1-2 days) ventricular and atrial cardiomyocytes were prepared as described.[27] After 24-hr attachment of primary myocytes in 6-well plates, at a density of 0.8×10^6 cells per well, cells were transfected following the calcium phosphate method and cultured in 4% fetal calf serum maintenance medium with or without agents and incubated for 48-hr in a 10% CO_2 incubator. Then cells were harvested for luciferase and β-galactosidase assays.[28] Luciferase activities were normalized to their corresponding β-galactosidase activities to correct for variation in transfection efficiency. Agonist LG153 (RXR specific) and TTNPB (RAR specific) were kindly provided by Dr. Richard Heyman (Ligand Pharmaceuticals, San Diego CA), BRL49653 was a gift from Dr. Robert Henry (VA Hospital, La Jolla CA). Clofibrate and all-trans RA were purchased from Sigma Chemicals (St. Louis MO) and Carboprostacyclin from Cayman Chemicals (Ann Arbor MI).

Gel shift: The gel mobility shift assays were performed as previously described.[29] The following oligonucleotides were used as probes and competitors (only the sense strand sequences are shown here): 5'-GTGAGGAAGGTCACAGGTCA-TAGTCCTGCA-3' (cDR-1); 5'-GTGAGGATGAACACTGAACCTAGTCAG-3' (mlcDR-1). For the competition studies, a 10 fold molar excess of unlabeled competitor was used, unless otherwise specified in the figure legends. For the gel shift mobility shift assays 0.2 μg of GST-purified nuclear receptors were incubated with 3μl of 5x binding buffer (50 mM HEPES pH 7.9; 375 mM KCl; 5 mM DTT; 5mM EDTA and 50% v/v glycerol), 1.5 μg poly (dI-dC), 2 μl LG153 ($5x10^{-7}$M). Incubations were performed for 10 min. on ice. DNA fragments were labeled using T4 polynucleotide kinase with [γ-^{32}P]ATP. To each sample 20,000cpm of labeled probe was added.

Data analysis

Data are represented as relative luciferase values. The data are based on three to six independent myocyte isolation and transfection experiments. For statistical analysis the ANOVA test was used. P value < 0.05 was considered statistically significant.

Results

The MLC-2a gene is persistently expressed in the ventricles of the RXRα mutant embryo, providing a model of defects in the maturation of ventricular muscle cell lineages.[22] To analyze the mechanisms by which retinoids regulate the expression of cardiac chamber-specific markers, we cloned the promoter of the MLC-2a gene and studied its transcriptional potential in various cell types. A 1 Kb fragment of the MLC-2a promoter fused to a luciferase reporter gene (1MLC-2a/Luc) displayed similar levels of activity in atrial and ventricular cells (figure 1), while no significant luciferase over the background was detected in COS cells, thus suggesting that this promoter fragment behaves cardiac specific, but does not distinguish atrial from ventricular cells.

MLC-2a promoter contains active retinoid response elements.

Since deficiency of RXRα is associated with impairment of MLC-2a downregulation in ventricular myocytes[22], we sought to determine if this reflected a direct effect of retinoids on MLC-2a transcription by transiently transfecting the 1MLC-2a/Luc reporter gene into neonatal rat ventricular cardiomyocytes. As displayed in figure 2, 1MLC-2a/Luc was down-regulated after treatment with the RXR-specific agonist LG153. The decrease in reporter activity was dependent on the dose

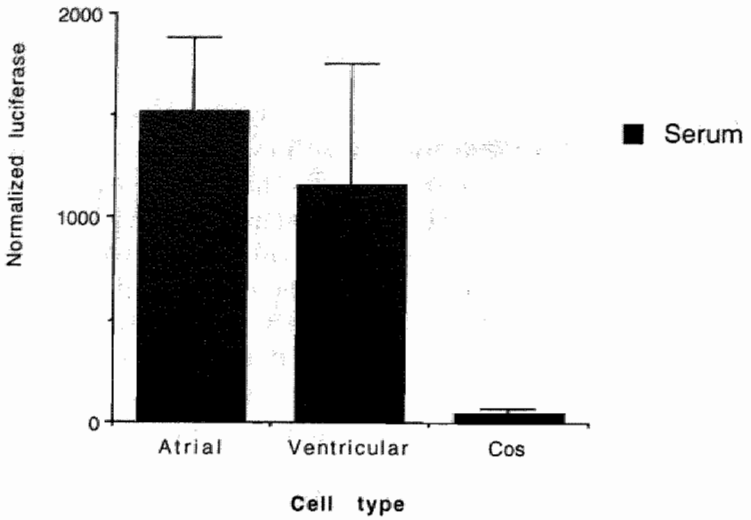


Figure 1. Tissue specificity of 1Kb MLC-2a promoter. Atrial and ventricular myocytes as well as COS cells were transfected with the 1MLC-2a luciferase reporter gene construct. Luciferase activity is shown as arbitrary units normalized to CMV/β-galactosidase activity.

of agonist, with a 50% repression (IC50) at 10^{-6} M, a retinoid concentration within physiological levels and comparable to dose-response profile for transcriptional activation mediated by retinoids.[30] We did not detect any significant response of the reporter gene level to treatment with all-trans-retinoic acid (atRA) the natural ligand for RAR (figure 2A) or with a stabilized RAR-specific agonist (TTNPB).

We determined whether overexpression in ventricular cells of RAR, RXR or a dominant negative RAR403 receptor had any additional effect on MLC-2a promoter activity via cotransfection studies. Overexpression of RAR did not alter 1MLC-2a/Luc basal activity and resulted in no significant response to retinoid treatment. Cotransfection of RXRα did not modify 1MLC-2a/Luc basal levels but an enhanced repressor effect was detected after incubation with 5×10^{-7} LG153 (figure 2B). No effect was observed after atRA (5×10^{-7}) treatment concentration (data not shown). Cotransfected 1MLC-2a/Luc with an engineered RAR dominant negative receptor (RAR403)[31], displayed an increased basal activity of the reporter gene and blunted response to retinoid agonists (figure 2b). These results document that the neonatal rat ventricular cells contain functional retinoid receptors that can act to directly repress the MLC-2a promoter and that ligand-mediated activation of RXR is required for MLC-2a downregulation.

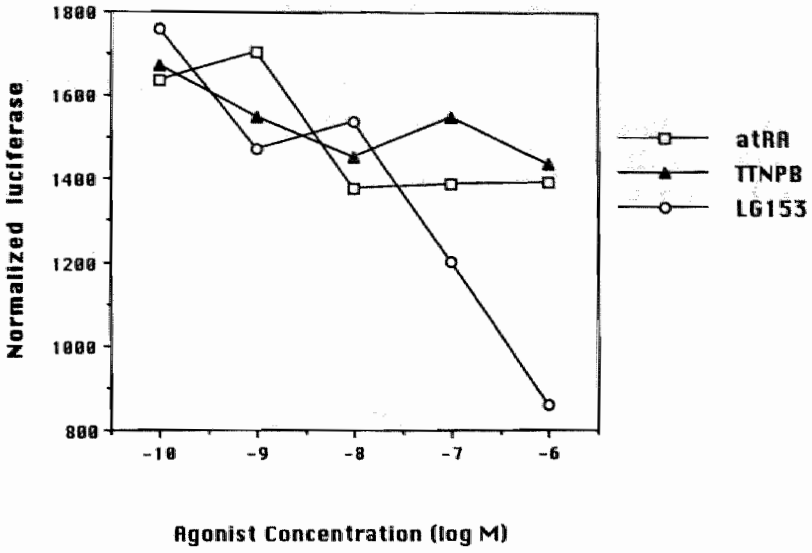


Figure 2A. Dose-response curve of MLC-2a-luciferase reporter gene to treatment with retinoid agonists. Dose-response of a 1Kb promoter fragment fused to a luciferase reporter gene (1MLC-2a-Luc) to the agonists: all-trans retinoic acid (atRA) and the synthetic agonists LG153 (RXR specific) and TTNPB (RAR specific).

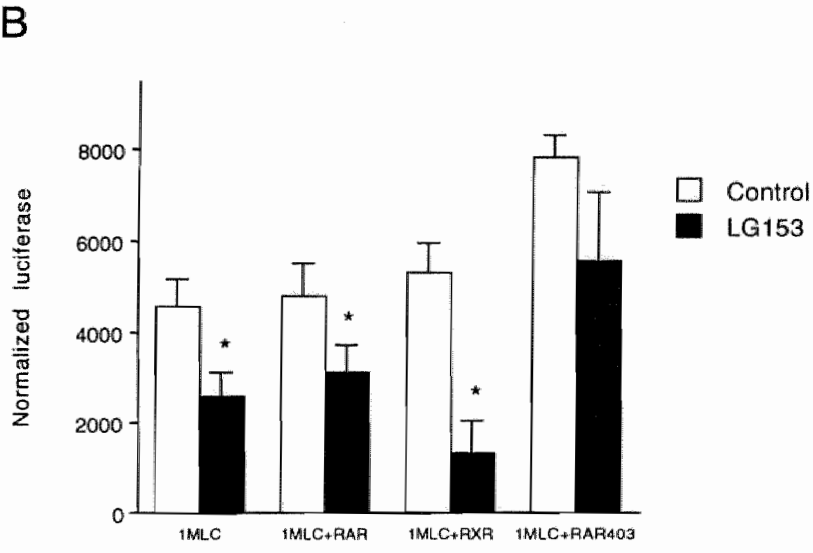


Figure 2B. 1MLC-2a/Luc was cotransfected into ventricular myocytes in the presence or absence of plasmids that contained human RARα RXRα or the dominant negative RAR (RAR403) driven by a CMV promoter. LG153 ($5x10^{-7}$M) was added to the culture media for 48hr. Bars represent luciferase activity normalized to CMV/β-galactosidase activity (* p value of <0.05, ANOVA test).

Retinoid-induced repression maps to a DR-1 element.

The MLC-2a promoter fragment contains in its sequences five putative hormone responsive elements (HRE). One estrogen receptor palindromic element is positioned at - 1622 bp. The 1MLC-2a/luc construct contains the remaining four HRE, which are Direct Repeats (DR) and conserved RAREs. These RAREs consist of one DR-1 and three DR-5 elements. The DR-1 and the most distal DR-5 element are oriented towards the 5' end of the gene, while the most proximal DR-5s are facing towards the transcription initiation site (figure 3D). Small divergences of these RAREs relative to the canonical consensus element were found. The DR-1 had one base pair substitution while the putative DR-5 elements displayed a higher degree of divergence with one or two substitutions. As described above, downregulation of 1MLC-2a/Luc reporter gene was selectively achieved after treatment with moderate doses of RXR-specific agonist.

Deletion of 0.1 Kb at the 5' end resulted in removal of the DR 1 element (0.93MLC-2a/Luc). This construct displayed a basal activity comparable to 1MLC-2a/Luc. However, 0.93MLC-2a/Luc had a blunted response to LG153 treatment (figure 3A). The generation of a reporter construct (mt1MLC-2a/Luc) in which the DR-1 element was substituted by a Pst I restriction site, in the context of the 1 Kb MLC-2a promoter provided additional support to the role of mlcDR-1 in the regulation of MLC-2a expression in response to retinoids. As noted in studies with the 0.93MLC-2a/Luc reporter gene, mt1MLC-2a/Luc displayed an overall basal activity comparable to its wild type 1 Kb counterpart and a blunted response to the RXR-specific agonist LG153 and no significant response was observed after atRA treatment (figure 3A). These results suggested that the mutated DR-1 element in the MLC-2a promoter (mlcDR-1) served as the major regulator of retinoid-induced repression of the MLC-2a gene transcription. In addition, retinoid-mediated repression appeared specific for mlcDR-1, since reporter vectors containing either DR1 (CRBP-II)[25]or DR-5 (βRE-1)[26] consensus elements were not down-regulated upon ligand incubation (figure 3B).

To determine whether mlcDR-1-mediated repression was dependent on its promoter context, we cloned one copy of mlcDR-1 containing nucleotides -951 to -927 upstream of the thymidine kinase (tk) basal promoter in the pT109 vector,[23] to generate a DR1pT109 construct. Addition of LG153 (RXR-specific agonist) to the culture media did not significantly inhibit DR1pT109 reporter expression. Likewise, no effect was observed after treatment with all-trans-retinoic acid (figure 3C) or the RAR-specific agonist TTNBP. Cotransfection of a human RARα expression plasmid did not modify reporter expression. However, cotransfection experiments with DR1pT109 and a RXR expression vector down-regulated reporter activity after addition of either LG153 or atRA to the culture media (figure 3C). The repressor effect observed after addition of atRA to the media could be explained by a certain degree of intracellular conversion of atRA

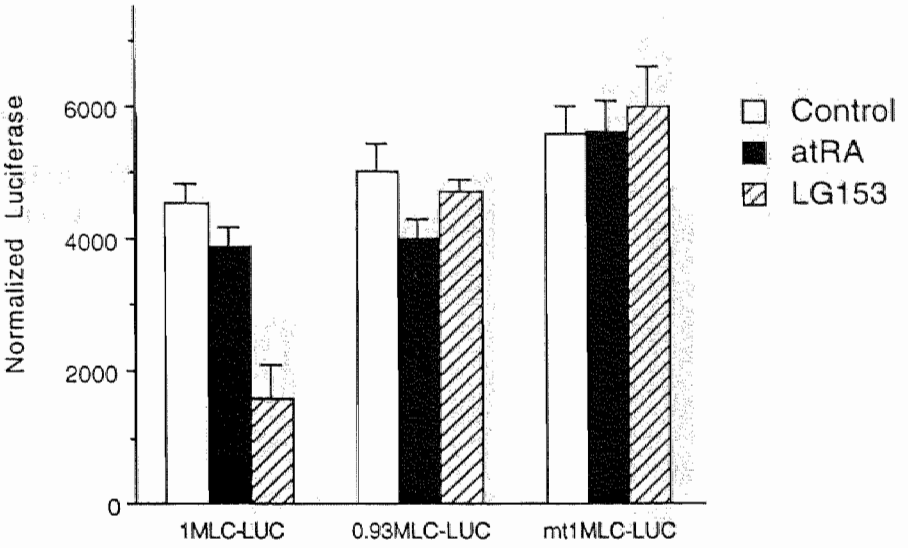


Figure 3A. Analysis of the activity of mlcDR-1. Reporter plasmids containing: 1Kb-luciferase (1MLC-2a/Luc), a deletion of DR-1, 0.93Kb-luciferase (0.9MLC-2a/Luc) or the 1Kb promoter with a mutated DR-1 element (mt1MLC-2a/Luc) were separately transfected into ventricular myocytes. Cells were incubated with all-trans retinoic acid (atRA, $5x10^{-7}$M) or LG153 ($5x10^{-7}$M) respectively for 48 hrs prior to determination of luciferase activity.

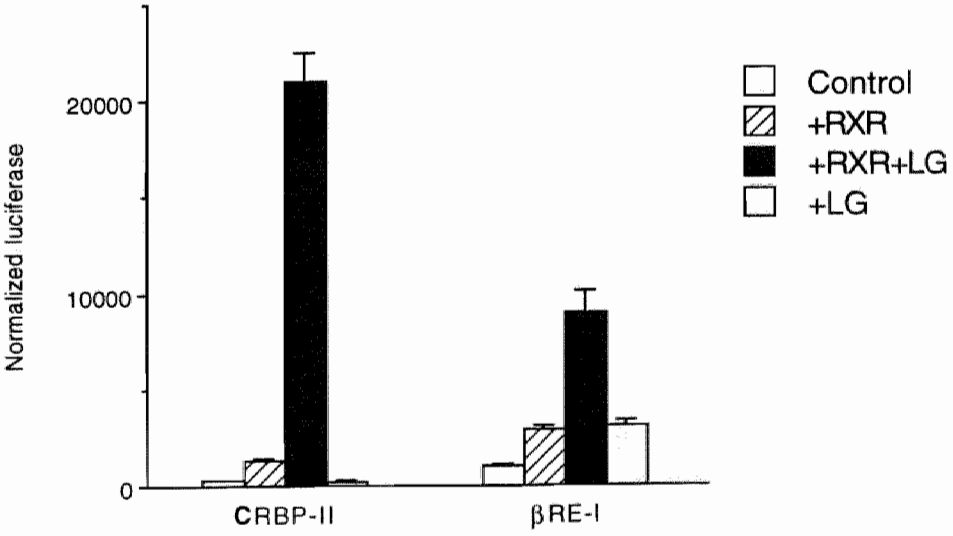


Figure 3B. Response of consensus DR-1 (CRBP-II) and DR-5 (βRE-I) reporter constructs to 48 hrs of treatment with LG153(10^{-7}MLG) in cultured and RXR transfected ventricular myocytes. Bars represent luciferase reporter activity normalized to a CMV/β-galactosidase control.

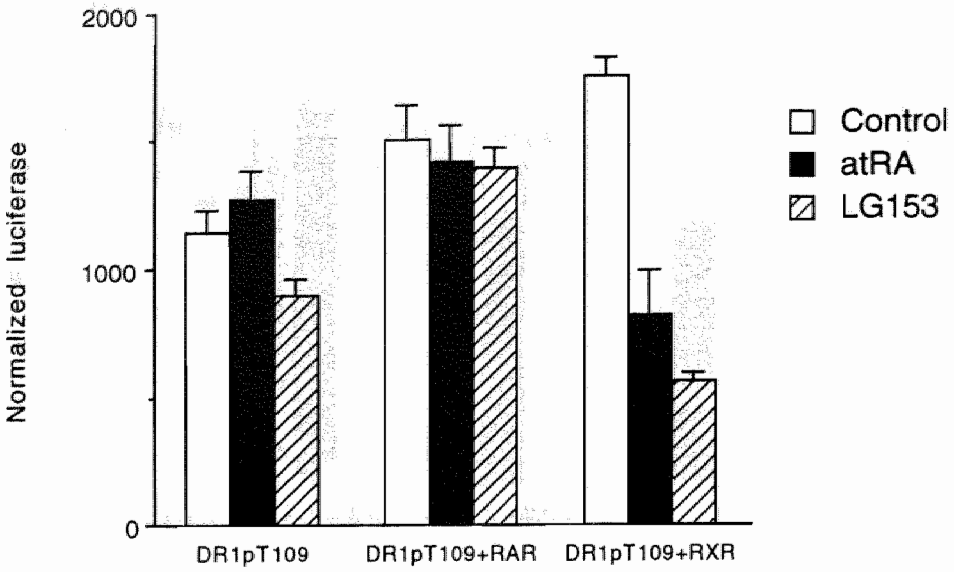


Figure 3C. A single DR-1 element upstream of a thymidine kinase promoter-luciferase reporter (DR1pT109) was transfected into ventricular myocytes with or without cotransfection of plasmids containing human RXRα or RARα driven by a CMV promoter.

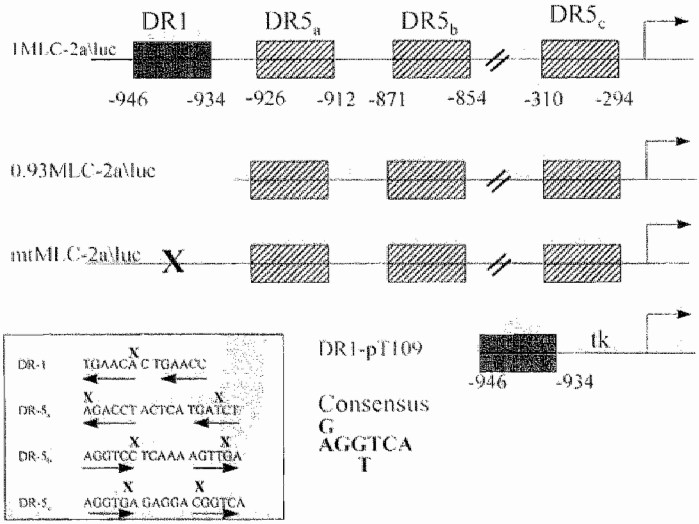


Figure 3D. Distribution of the putative retinoid response elements in a 2Kb fragment ot the atrial myosin light chain-2 (MLC-2a) promoter. The sequence and orientation of these element as well as the consensus sequence are indicated.

into 9-cis RA. Thus, in the context of a heterologous promoter, a repressor effect is only observed after overexpression and activation of RXR. This is in contrast with our observations in the full 1 Kb MLC-2a promoter, where activation of the endogenous RXR receptors is enough to achieve a significant repressor effect. These results indicate that mlcDR-1 may be dependent on its promoter context, since the isolated element requires higher doses of ligand-activated RXR to function as a transcriptional repressor element.

mlcDR-1 discriminates the recruitment of nuclear receptors

RXRα is a common factor in the different combination of nuclear receptors that confer transcriptional activation or repression upon binding to DR-1 elements. Thus, DR-1 elements can bind either RXR homodimers or heterodimeric combinations of RXR with RAR (retinoic acid receptor) or PPAR (peroxisome proliferator-activated receptor).[12] DR-1 elements act as potent repressor sites in combination with RAR by suppressing RXR ligand binding.[32] However, mlcDR-1 acts as a repressor only after ligand-mediated activation suggesting that its interaction with the nuclear receptors may be different from that of the consensus DR-1 site.[25,33] To test this hypothesis, we performed gel mobility shift assays either with mlcDR-1 or a consensus DR-1 probe (cDR-1), using purified GST-RXRα, GST-RARα and GST-PPARγ fusion proteins in the presence of specific RXR ligands. In agreement with previous reports [25,33,34], cDR-1 bound to RXR/RXR, RXR/RAR and RXR/PPAR dimers and was specifically competed by addition of a 10 fold excess of cold cDR-1. In contrast, a 10 fold excess of cold mlcDR-1 displayed poor competition for RXR/RXR and RXR/RAR but competition for the RXR/PPAR complex (figure 4). Similarly, mclDR-1 bound either RXR/RXR, RXR/RAR and RXR/PPAR complexes (not shown). Thus, mclDR-1 behaves as a retinoid receptor binding site and contains in its sequence enough information to discriminate the recruitment of specific nuclear receptors and suggest that mlcDR-1 acts as a preferential site for RXR/PPAR whose activation would decrease the transcriptional level of the MLC-2a promoter.

Cooperative effect of RXR and PPAR in MLC-2a transcriptional regulation

The above results suggest that RXR and PPARγ could be partners responsible for MLC-2a downregulation, however, sofar experiments with overexpression of PPARγ did not result in downregulation of MLC-2a in ventricular cells (data not shown). Since MLC-2a transcriptional repression in ventricular cells is a dynamic process that starts around embryonic day 9 and is complete at day 14 in mouse,[35] the possibility exists that cotransfection studies in ventricular cells were masked by an already repressed MLC-2a promoter.

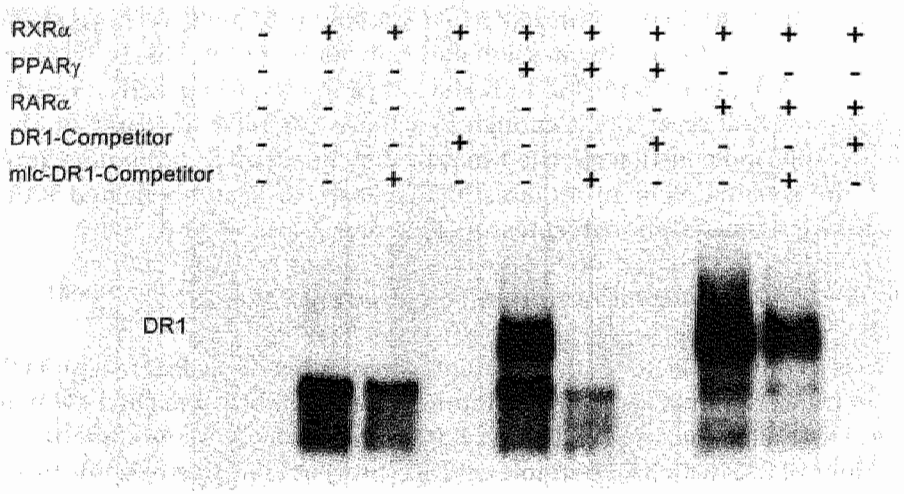


Figure 4. Gel mobility shift assay. For the characterization of nuclear receptors that interact with the DR-1 element. GST purified nucleoproteins as indicated (RARγ, RXRα and RAR 403) were incubated with cDR-1 probe labeled with γ^{32}P-ATP. Competition was performed with either cold cDR-1 or cold mlc-DR-1 respectively. The bands shown differ in size depending on the protein-DNA complex formed. Partial competition is shown after adding mlcDR-1 competitor in the presence of RXRα and PPARγ (lane 6). Cold cDR1 results in full competition as expected (lane 4, 7, 10).

To approach this problem and analyze further the mechanism of MLC-2a regulation we performed cotransfection studies on isolated atrial myocytes followed by incubation with RXR and PPAR-specific agonists. Similarly to the results in ventricular myocytes, the activation of RXR by LG153 resulted in downregulation of the 1MLC-2a/Luc reporter gene. Cotransfection of expression vectors for PPARα, PPARγ and PPARδ did not significantly modify MLC-2a reporter expression. However, the addition of either PPARα or PPARγ specific-agonists displayed a of 80% and 50% repression of transcriptional activity respectively. No repression was observed upon the addition of the PPARδ-specific agonist (figure 5). These results suggest that PPARα or PPARγ are the endogenous partner of RXRα for MLC-2a repression in atrial cardiomyocytes.

Discussion

The data presented here, show that a MLC2a-luciferase reporter gene is downregulated by retinoids in ventricular myocytes after specific activation of RXR. This action is mediated through a DR-1 retinoid responsive element located at position -934 of the MLC-2a 5' flanking region. These results recapitulate the aberrant ventricular expression of the atrial myosin light chain (MLC-2a) in the

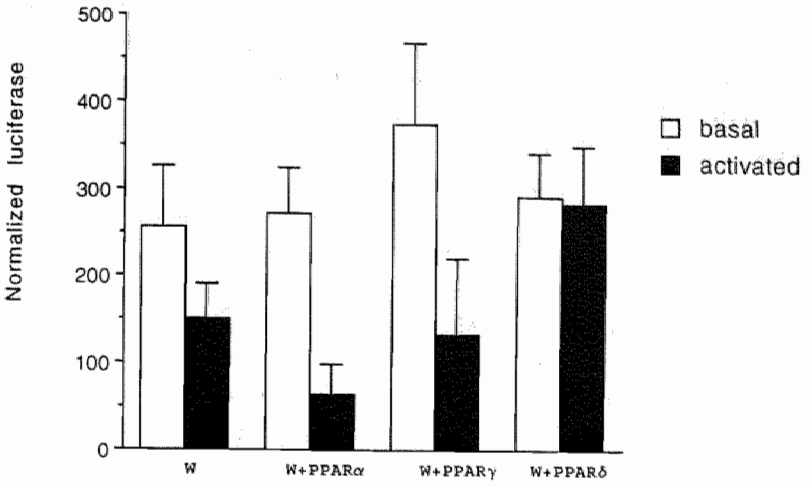


Figure 5. MLC-2a promoter in atrial cardiomyocytes. Atrial cardiomyocytes were transfected with wildtype (W) 1MLC-2a/luc, open bars and cotransfected with RXRα and various PPAR isoforms, in addition to agonist treatment (closed bars). W: treatment with LG153. W+PPARα: clofibrate, W+PPARγ: BRL, W+PPARδ: carboprostacyclin.

context of the RXRα null mutant mice and provide a likely explanation on the molecular mechanisms that regulate ventricular myocyte gene expression. Activation of RXR in DR-1 sites usually results in transcriptional activation that may be blocked by RAR. RAR prevents RXR from binding its ligand through allosteric interactions.[32] However, when RAR is occupied by its own specific ligand, RXR ligands are able to activate the RXR homodimer.[16,36] Because RAR/RXR heterodimers bind to DR-1 sites with higher affinity than RXR/RXR homodimers, escape from the inhibitory effects of RAR would require an enhanced expression of RXR to drive heterodimeric associations with other nuclear receptors (such as PPAR) that do not inhibit the response to RXR ligands.[32] The DR-1 element in the MLC-2a promoter has distinct properties, because it binds RXR/RXR and RXR/RAR with low affinity and responds to ligands with a decrease in transcriptional potential. Neither treatment with saturating doses of RAR agonists nor overexpression of RXR results in a relieved repression. In addition, stimulation of PPAR or RXR results in downregulation of reporter activity. Both deletion and site-directed mutation of the DR-1 site (mlcDR-1) in the MLC-2a promoter block the ligand-mediated repression, suggesting that mlcDR-1 is a key regulator of this effect.

The mlcDR-1 site behaved like a repressor in a heterologous promoter only at high doses of activated RXR, suggesting that mlcDR-1 is dependent on its promoter context and that the activation of RXR in mlcDR-1 may be recruiting a

cellular co-factor that would modify retinoid receptor interactions with the basal transcriptional machinery. In this model, RXR is more likely a key player than a silent partner, and activation of RXR seems to be an absolute requirement for the retinoid mediated repression of MLC-2a, in the ventricular chamber during development.

The differences in behavior between atrial and ventricular myocytes will be addressed by additional experiments, treating ventricular cardiomyocytes with various PPAR agonists, with and without overexpression of the PPA receptors.

The effect of retinoids in the heart has been extensively studied using the atrial natriuretic factor (ANF) expression as a marker for cardiac hypertrophy. ANF induction during hypertrophy is reversed after retinoid treatment.[24] Whether the effect of retinoids on the suppression of ANF induction is dependent on the ANF promoter is not clear. Vitamin D3 and RA antagonize endothelin-stimulated hypertrophy in neonatal rat cardiomyocytes[2] and vitamin D3 suppression of ANF transcription maps to a site located in -104 of the human ANF promoter.[37] In contrast, vitamin D3 does not block phenylephrine-induced hypertrophy.[24] These results suggest a key role for RXR as a regulator of cardiac hypertrophy in at least two different pathways and using alternate mechanisms: either as a direct transcriptional regulator [37] or blocking an intermediate step in the adrenergic transduction signaling.[24] Retinoid-mediated mechanisms for maintenance of the adult phenotype and for chamber specification during embryogenesis might then be different. This is in agreement with the fact that RAR activation is required for repression of hypertrophy[24], while RXR activation (our results) is required for ventricular downregulation of atrial MLC-2 expression. The ability of retinoids to suppress ventricular hypertrophy and to block ANF induction in an in vitro model of cardiac hypertrophy[24], suggest that retinoid signaling may have a critical role both during development and in the maintenance of the adult cardiac phenotype.

Conclusions

These results provide evidence for a direct role of RXR in the regulation of chamber-specific gene expression, as RXR and PPAR activation can downregulate MLC-2a promoter activity in cardiomyocytes. A retinoid responsive element on the MLC-2a promoter (DR-1) mediates the interaction with the RXR pathway and is responsible for exogenous MLC-2a promoter activity.

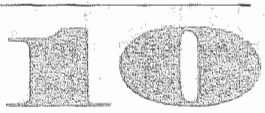
Literature

1. Wilson JG, Warkany J: Aortic-arch and cardiac abnormalities in the offspring of vitamin A deficient rats. *Am J Anat* 1949;85:113-155
2. Wu J, Garami M, Cheng T, Gardner DG: 1,25(OH)2 vitamin D3, and retinoic acid antagonize endothelin-stimulated hypertrophy of neonatal rat cardiac myocytes. *J* Clin Invest 1996;97:1577-1588
3. Evans RM: The steroid and thyroid hormone receptor superfamily: transcriptional regulators of development and physiology. *Science* 1988;240:889-895
4. Smith MA, Parkinson DR, Cheson BD, Friedman MA: Retinoids in cancer therapy. *J* Clin Oncol 1992;10:839-864
5. Gudas LJ, Sporn MB, Roberts AB: Cellular biology and biochemistry of retinoids, in Roberts AD, Goodman DS (eds): *The retinoids: biology, chemistry and medicine.* New York, Raven Press, 1994, pp 443-520
6. Beato M: Transcriptional control by nuclear receptors. *FASEB* 1991;5:2044-2051
7. Glass CK: Differential recognition of target genes by nuclear receptors monomers, dimers and heterodimers. *Endocrinol Rev* 1994;15:391-407
8. Umesono K, Murakami KK, Thompson CC, Evans RM: Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. *Cell* 1991;65:1255-1266
9. Naar AM, Boutin JM, Lipkin SM, Yu VC, Holloway JM, Glass CK, Rosenfeld MG: The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors. *Cell* 1991;65:1267-1279
10. Yu VC, Naar AM, Rosenfeld MG: Transcriptional regulation by the nuclear receptor superfamily. *Curr Opin Biotechnol* 1992;3:597-602
11. Parker MG: Steroid and related receptors. *Curr Opin Cell Biol* 1993;5:499-504
12. Mangelsdorf DJ, Evans RM: The RXR heterodimers and orphan receptors. *Cell* 1995;83:841-850
13. Schrader M, Nayeri S, Kahlen JP, Muller KM, Carlberg C: Natural vitamin D3 response elements formed by inverted palindromes: polarity-directed ligand sensitivity of vitamin D3 receptor-retinoid X receptor heterodimer-mediated transactivation. *Mol Cell Biol* 1995;15:1154-1161
14. Willy PJ, Umesono K, Ong ES, Evans RM, Heyman RA, Mangelsdorf DJ: LXR, a nuclear receptor that defines a distinct retinoid response pathway. *Genes Dev* 1995;9:1033-1045
15. Perlmann T, Jansson L: A novel pathway for vitamin A signaling mediated by RXR heterodimerization with NGFI-B and NURR1. *Genes Dev* 1995;9:769-782
16. Forman BM, Umesono K, Chen J, Evans RM: Unique response pathways are established by allosteric interactions among nuclear hormone receptors. *Cell* 1995;81:541-550

17. Issemann I, Prince RA, Tugwood JD, Green S: The peroxisome proliferator-activated receptor:retinoid X receptor heterodimer is activated by fatty acids and fibrate hypolipidaemic drugs. *J Mol Endocrinol* 1993;11:37-47
18. Kastner P, Mark M, Chambon P: Nonsteroid nuclear receptors: what are genetic studies telling us about their role in real life? *Cell* 1995;83:859-869
19. Beato M, Herrlich P, Schutz G: Steroid hormone receptors: many actors in search of a plot. *Cell* 1995;83:851-857
20. Gruber PJ, Kubalak SW, Pexieder T, Sucov HM, Evans RM, Chien KR: RXR alpha deficiency confers genetic susceptibility for aortic sac, conotruncal, atrioventricular cushion, and ventricular muscle defects in mice. *J Clin Invest* 1996;98:1332-1343
21. Sucov HM, Dyson E, Gumeringer CL, Price J, Chien KR, Evans RM: RXRa mutant mice establish a genetic basis for vitamin A signaling in heart morphogenesis. *Genes Dev* 1994;8:1007-1018
22. Dyson E, Sucov HM, Kubalak SW, Schmid Schonbein GW, DeLano FA, Evans RM, Ross J, Jr., Chien KR: Atrial-like phenotype is associated with embryonic ventricular failure in retinoid X receptor alpha -/- mice. *Proc Natl Acad Sci U S A* 1995;92:7386-7390
23. Nordeen SK: Luciferase reporter gene vectors for analysis of promoter and enhancers. *Biotech* 1988;6:454-457
24. Zhou MD, Sucov HM, Evans RM, Chien KR: Retinoid-dependent pathways suppress myocardial cell hypertrophy. *Proc Natl Acad Sci U S A* 1995;92:7391-7395
25. Mangelsdorf DJ, Umesono K, Kliewer SA, Borgmeyer U, Ong ES, Evans RM: A direct repeat in the cellular retinol binding protein type II gene confers differential regulation by RXR and RAR. *Cell* 1991;66:555-561
26. Sucov HM, Murakami KK, Evans RM: Characterization of an autoregulated response element in the mouse retinoic acid receptor type b gene. *Proc Natl Acad Sci U S A* 1990;87:5392-5398
27. Knowlton KU, Baracchini E, Ross RS, Harris AN, Henderson SA, Evans SM, Glembotski CC, Chien KR: Co-regulation of the atrial natriuretic factor and cardiac myosin light chain-2 genes during alpha-adrenergic stimulation of neonatal rat ventricular cells. Identification of cis sequences within an embryonic and a constitutive contractile protein gene which mediate inducible expression. *J Biol Chem* 1991;266:7759-7768
28. Thorburn A, Thorburn J, Chen SY, Powers S, Shubeita HE, Feramisco JR, Chien KR: HRas-dependent pathways can activate morphological and genetic markers of cardiac muscle cell hypertrophy. *J Biol Chem* 1993;268:2244-2249
29. Navankasattusas S, Zhu H, Garcia AV, Evans SM, Chien KR: A ubiquitous factor (HF-1a) and a distinct muscle factor (HF-1b/MEF-2) form an E-box-independent pathway for cardiac muscle gene expression. *Mol Cell Biol* 1992;12:1469-1479

30. Rohrer DK, Hartong R, Dillmann WH: Influence of thyroid hormone and retinoic acid on slow sarcoplasmatic reticulum Ca (2+) ATPase and myosin heavy chain alpha gene expression in cardiac myocytes: delineation of cis active DNA elements that confer responsiveness to thyroid hormone but not to retinoic acid. *J Biol Chem* 1991;266:8638-8646
31. Damm K, Heyman RA, Umesono K, Evans RM: Functional inhibition of retinoic acid response by dominant negative retinoic acid receptor mutants. *Proc Natl Acad Sci U S A* 1993;90:2989-2993
32. Kurokawa R, DiRenzo J, Boehm M, Sugarman J, Gloss B, Rosenfeld MG, Heyman RA, Glass CK: Regulation of retinoid signalling by receptor polarity and allosteric control of ligand binding. *Nature* 1994;371:528-531
33. Mangelsdorf DJ, Umesono K, Evans RM: The retinoid receptors, in Goodman DS, Roberts AD (eds): *The retinoids: Biology, chemistry and medicine.* New York, Raven Press, 1994, pp 319-349
34. Durand B, Saunders M, Leroy P, Leid M, Chambon P: All-trans and 9-cis retinoic acid induction of CRABPII transcription is mediated by RAR-RXR heterodimers bound to DR1 and DR2 repeated motifs. *Cell* 1992;71:73-85
35. Kubalak SW, Miller-Hance W, O'Brien T, Dyson E, Chien K: Chamber-specification of atrial myosin light chain-2 expression precedes septation during mouse cardiogenesis. *J Biol Chem* 1994;269:16961-16970
36. Roy B, Taneja R, Chambon P: Synergistic activation of retinoic acid (RA)-responsive genes and induction of embryonal carcinoma cell differentiation by an RA receptor alpha (RAR alpha)-, RAR beta-, or RAR gamma-selective ligand in combination with a retinoid X receptor-specific ligand. *Mol Cell Biol* 1995;15:6481-6487
37. Li Q, Gardner DG: Negative regulation of the human atrial natriuretic peptide gene by 1,25-dihydroxyvitamin D3. *J Biol Chem* 1994;269:4934-4939

Chapter



General discussion

In this thesis various aspects of the regulatory myosin light chain (MLC-2) genes were investigated. The atrial and ventricular light chain proteins have been used to phenotype cardiac myocytes derived from an *in vitro* embryonic system. The functional studies performed on ion channels, support the value of this *in vitro* system to study early events in cardiogenesis (chapter 5). Not only the onset of functional ion channel expression is comparable, also the current characteristics and the changes in current modulation are similar in fetal cardiomyocytes compared with *in vitro* embryoid body derived cardiomyocytes. The cardiomyocytes isolated from embryoid bodies are a mixture of undifferentiated, atrial, and ventricular cardiomyocytes. Therefore these cardiomyocytes can not replace the rat neonatal cardiomyocyte cell culture for transient transfections. However use of embryonic stem cells for gene targeting approaches and embryoid body formation can be used to study genetic modifications, that appear to be embryonic lethal. For some questions concerning ion channel formation the embryoid body system can provide the answers, without generating new mouse lines. New techniques like a tetraploid system for full embryonic development from a genetically modified stable embryonic stem cell line is one of the recent developments.[1,2]

Cardiac gene transcription is a fascinating and complex field where new key players on the protein and DNA level come in focus on a regular bases. Any review is outdated before it is published and can not nearly be complete. The review in this thesis provides insight in the various mechanisms involved in transcription regulation. The methods to identify new DNA\RNA binding proteins are discussed separately. The methodology outlined in chapter 3 was used in the work described in following chapters and led for instance to the cloning of mSAP49, a splice factor with an interesting expression pattern during cardiac development. Many facts about precursor RNA splicing are still unknown, and a potential role for non-coding splice products (introns) is a major topic in ongoing transcription regulation research. Although splicing can be an automatic event in many genes the finding of tissue specific alternative splicing, indicates the importance of additional splice factors. Splicing events common in one cell type, do not take place in other cell types. The functional role of mSAP49 is yet unknown as well as its target genes. We have shown binding of mSAP49 to the inhibiting HF3 *cis* element of the MLC-2v promoter. However whether this is an *in vitro* finding or an *in vivo* relevant DNA-protein interaction will have to be addressed in mSAP49 knock out studies.

A major part of the thesis is aimed at the identification and characterization of the MLC-2a promoter. The work presented has focused on the cell type specificity. During *in vitro* studies the regulatory DNA appeared to be cardiac specific. However we were not able to show atrial myocyte specific expression using the *in vitro* system of primary cultures of neonatal rat atrial and ventricular myo-cytes. Potentially this could be explained by the absence of an atrial specific

enhancer in our constructs. However the entire 5' flanking region of the MLC-2a gene spanning the upstream glucokinase gene to the first exon of the MLC-2a gene, was studied. Furthermore, the potential role of the first intron in transcription regulation was analyzed (transient transfection studies). In addition the 3' flanking region was sequenced to identify additional *cis* elements. In addition we compared regulatory sequences of MLC-2a with the ANF gene. So far we were not able to locate potential atrial specific enhancers. From literature we know that tissue specific enhancers can be positioned at a large distance from the essential promoter and coding region. For instance, distances up to 40 Kb have been reported for the α-globin gene.[3] Obviously the value of the *in vitro* cardiomyocyte studies is limited due to the finding of re-expression of the endogenous MLC-2a gene in ventricular myocytes in culture. The answer of atrial specificity will have to come from *in vivo* evaluation through the generation of transgenic mice harboring MLC-2a promoter fragments linked to the β-galactosidase reporter gene and *in vitro* from the tetraploid embryonic system. The importance of the initiator *cis* element for initiation of transcription of the MLC-2a gene has been illustrated. In a separate chapter the interaction of the MLC-2a regulatory DNA with several members of the superfamily of nuclear receptors has been studied. The results described here could explain the persistent expression of the MLC-2a gene in ventricular tissue of homozygous RXR-α knock out embryos. However the issue remains as to how the RA signaling pathway is influenced by the *in vitro* culture conditions, to which the primary cultures of myocytes are subjected.

Apparently the MLC-2a promoter can be stimulated by another family member of the nuclear receptor belonging to the superfamily namely: the estrogen receptor. Treatment of cardiomyocytes with estrogen showed a clear induction of promoter activity. The role of the estrogen receptor in mediating this response will be studied, by gel mobility shift assay and site directed mutagenesis. In addition to estrogen also phenylephrine and endothelin-1 activate the MLC-2a promoter, potentially involving AT rich *cis* elements. In addition the re-expression of the MLC-2a gene under hypertrophic conditions was documented. Myocyte treatment with neurohumoral factors stimulated endogenous MLC-2a expression. Suggesting a similar behavior of the isolated promoter fragment to the endogenous MLC-2a promoter.

In ventricular tissue of patients with dilated cardiomyopathy, the re-expression of MLC-2a was shown at the mRNA and protein level. The cause of the disease in these patients is unknown and can be different in all patients suffering from this disease. Therefore conclusions should be made very carefully. The results shown here could be indicative for ongoing cardiomyocyte hypertrophy, which in the selected patients is insufficient to prevent dilatation of the left ventricular chamber. The re-expression of ANF and MLC-2a are probably signs of the hypertrophic response.

Thanks to molecular genetics the list of genetic causes for inherited cardiac diseases is growing steadily.[4,5] However the molecular defects leading to primary dilated cardiomyopathy are still obscure. Although 3 chromosomal loci have been identified by haplotyping (1p1,[6] 3p25[7], 9q13[8]), no causative genes are known today. Mutations in the myosin light chains are a cause for hypertrophic cardiomyopathy,[9] however no mutations were found in patients presenting with a dilated cardiomyopathy.[10] In the patients included, the cause of the cardiomyopathy is unknown and in all likelihood will vary for each patient. The molecular analysis presented in chapter 8 on human ventricular tissue samples will help to dissolve the signaling pathways and molecular mechanisms causing idiopathic dilated cardiomyopathy, also in individual non-familiar cases.

New information is always raising more questions then providing answers. This thesis is no exception to this rule. The data presented here mostly provide tools for cardiovascular research in various fields. One of the interesting facts about the MLC-2a gene is the very early expression of the gene in cardiogenesis. Moreover in various transcription factor KO experiments with severely impaired cardiac development the MLC-2a gene is expressed at the expected time point and localization. Expression of the MLC-2a gene is independent of the reduced transcription factor levels, in these experiments.[11-13] Apparently the transcription factors required for expression of this gene are expressed early. Up to now we do not know which factors are crucial for the first steps in cardiogenesis. Surely the MLC-2a promoter will help to elucidate the early molecular steps involved in the transition of cardiac precursor cell to cardiomyocyte in the near future.

Potentially three different promoters could be used in transgenic experiments to drive atrial specific gene expression. First of all the ANF promoter has been characterized and used successfully in transgenic experiments.[14,15] The most important limitation of the ANF promoter is the continuous activity in ventricular myocytes throughout development. The promoter is silenced shortly before birth. An alternative would be the MLC-1a promoter however, MLC-1a is expressed initially in the whole heart and skeletal muscle, and remains active after birth in Purkinje fibers. The downregulation of the MLC-1a promoter in the ventricle is seen around day 16 of murine embryonic development. Therefore the MLC-2a promoter has unique features. The promoter is activated early (day 6), cardiac specific during all developmental stages and only briefly expressed in ventricular tissue (day 8-12). However it is important to realize that the three promoters can be re-activated in ventricular tissue under hypertrophic conditions. The value of the various promoter fragments will have to be determined *in vivo*.

Implications and future research

Function is the key issue in modern molecular cardiology. To determine the mSAP49 function, one approach will be to knock out the gene through homologous recombination. By back crossing the heterozygous off-spring, homozygous knock out mice can be obtained, that are deficient in mSAP49. The cardiac phenotype of these mice will be studied, focussing primarily on congenital abnormalities, cardiac function and contractile protein expression.
The function of the MLC-2a promoter has to be studied *in vivo* before the next series of experiments can be initiated. Constructs with the MLC-2a promoter driving β-galactosidase have been made and were injected into oocytes to this purpose already. In addition experiments to study tissue specificity of the MLC-2a promoter in the tetraploid system have been initiated. For these studies stable embryonic stem cell lines are being made with the MLC-2a promoter\β-galactosidase transgene. These stem cells can be used to generate full grown mouse embryos in a tetraploid environment without going trough the breeding of mouse lines. The embryos can be stained *in toto* to study the transgene expression pattern.

The promoter potentially could be used to drive atrial specific transgene expression. However as mentioned before, under pathologic (hypertrophic) conditions the promoter might be activated in ventricular tissue. In addition the promoter can be used to drive recombinase expression to generate atrial specific KO lines. Ultimately these tools might enable studies on atrial myocyte action potential duration in relation to ion channel gene expression, as we can modify the expression levels of ion channels and record atrial myocyte electrical activity. To facilitate these future studies cellular, *ex vivo* and *in vivo* electrophysiology in mouse will need to be developed.

Molecular cardiology can be applied to analyze endocardial biopsies and tissue sections of cardiomyopathy patients. These studies will help to bridge the gap between basic science and clinical cardiology. Different hypertrophic stimuli will activate alternative signaling pathways leading to variation in the molecular response of cardiomyocytes. The re-expression of MLC-2a in cardiac tissue of cardiomyopathy patients is the final outcome of the activation of a signaling pathway. The deciphering of signaling pathways involved, may help us to identify crucial molecular switches in the development of the hypertrophic response, with potential therapeutic implications. The final goal of all the efforts to increase our knowledge on physiologic and pathologic molecular changes is to understand human cardiac development, physiology and disease.

Research in molecular cardiology is time consuming and expensive. Small contributions to the analysis of molecular events require a wide range of molecular tools and hard work. Fortunately, many researchers around the globe are aware of the difficulties of this line of research. Exchange of ideas and material

between different research groups is required to resolve the molecular puzzle. Molecular biology is now used for the development of new therapeutic drugs. However the direct use of genetic material for therapeutic approaches is hampered by the limitations of gene transfer into mammalian cells, especially in terminally dif-ferentiated cardiac muscle cells. Nevertheless, studies as presented in this thesis may result in new drugs to develop a tailored therapy for atrial disease and possibly for cardiomyopathy patients. Crucial in these therapeutic approaches will be the development of safe and secure ways of gene transfer.

Literature

1. Nagy A, Rossant J: Production of completely ES-cell derived fetuses, in Joyner AL (ed): *Gene targeting: a practical approach.* Oxford University Press, 1993, pp 147-180
2. Nagy A, Rossant J: Targeted mutagenesis: analysis of phenotype without germ line transmission. *J Clin Invest* 1996;97:1360-1365
3. Strauss EC, Andrews NC, Higgs DR, Orkin SH: In vivo footprinting of the human alpha-globin locus upstream regulatory element by guanine and adenine ligation-mediated polymerase chain reaction. *Mol Cell Biol* 1992;12:2135-2142
4. Meijer H, Doevendans PA, van Dantzig J, Schaap C, Jongbloed R, Geraedts J: On the clinical and genetical identification of familial cardiomyopathies. *Eur Heart J* 1997;Submitted:
5. Doevendans PA, van Dantzig J, Meijer H, Schaap C: Molecular genetics of human cardiomyopathies, in Peters R, Piek J (eds): *Molecular cardiology in clinical perspective.* Amsterdam, Knoll, 1997, pp 33-53
6. Kass S, MacRae C, Graber HL, Sparks EA, McNamara D, Boudoulas H, Basson CT, Baker PB, Cody RJ, Fishman MC, et al: A gene defect that causes conduction system disease and dilated cardiomyopathy maps to chromosome 1p1-1q1. *Nat Genet* 1994;7:546-551
7. Olson TM, Keating MT: Mapping a cardiomyopathy locus to chromosome 3p22-p25. *J Clin Invest* 1996;97:528-532
8. Krajinovic M, Pinamonti B, Sinagra G, Vatta M, Severini GM, Milasin J, Falaschi A, Camerini F, Giacca M, Mestroni L: Linkage of familial dilated cardiomyopathy to chromosome 9. Heart Muscle Disease Study Group. *Am J Hum Genet* 1995;57:846-852
9. Poetter K, Jiang H, Hassanzadeh S, Master SR, Chang A, Dalakas MC, Rayment I, Sellers JR, Fananapazir L, Epstein ND: Mutations in either the essential or regulatory light chains of myosin are associated with a rare myopathy in human heart and skeletal muscle. *Nat Genet* 1996;13:63-69

10. Holt JC, Caulfield JB, Norton P, Chantler PD, Slayter HS, Margossian SS: Human cardiac myosin light chains: sequence comparisons between myosin LC1 and LC2 from normal and idiopathic dilated cardiomyopathic hearts. *Mol Cell Biochem* 1995;145:89-96
11. Dyson E, Sucov HM, Kubalak SW, Schmid Schonbein GW, DeLano FA, Evans RM, Ross J, Jr., Chien KR: Atrial-like phenotype is associated with embryonic ventricular failure in retinoid X receptor alpha -/- mice. *Proc Natl Acad Sci U S A* 1995;92:7386-7390
12. Lyons I, Parsons LM, Hartley L, Li R, Andrews JE, Robb L, Harvey RP: Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeobox gene Nkx2-5. *Genes Dev* 1995;9:1654-1666
13. Lin Q, Schwarz J, Bucana C, Olson EN: Control of mouse cardiac morphogenesis and myogenesis by transcription factor MEF2C. *Science* 1997;276:1404-1408
14. Rockman HA, Ross RS, Harris AN, Knowlton KU, Steinhelper ME, Field LJ, Ross JJ, Chien KR: Segregation of atrial-specific and inducible expression of an atrial natriuretic factor transgene in an in vivo murine model of cardiac hypertrophy. *Proc Natl Acad Sci U S A* 1991;88:8277-8281
15. Field LJ: Atrial-natriuretic factor-SV40 large T antigen transgenes produce atrial tumors and cardiac arrhythmias in mice. *Science* 1988;239:10029-10033

Summary

In this thesis various aspects of the murine regulatory myosin light chain (MLC-2) genes were investigated. Two cardiac MLC-2 proteins have been identified, one is expressed in ventricular myocytes, while the other isoform is expressed in atrial myocytes. In chapter one a brief introduction into the nomenclature of myosin light chains , the sarcomeric stucture and molecular biology in general is provided. Preliminary data that provide the bases for this study are discussed together with an outline of work presented in this thesis.

The complexity of transcription regulation is highlighted in the second chapter. To provide some insight in gene transcription in the cardiovascular system the most important regulating systems are reviewed. In addition, previously identified transcription factors that have been shown to be important for gene transcription in the heart are listed, linked to the role of the various factors for specific contractile protein genes known to date. In chapter 3 the methods to study nucleic acid-transcription factor interaction are reviewed. Molecular biology has developed the tools to study the interaction of nuclear proteins with either DNA or RNA. The interaction can be analyzed at the basal level to identify the sequence of the nucleic acid bases required for binding or in functional assays to determine the role of the various transcription factors in cell culture and in the intact organism. The strategy to study new transcription factors has been applied to clone and characterize mouse SAP49. The human isoform of this protein was previously identified and appeared to be a RNA splice factor. The mSAP49 message exhibits an interesting expression pattern during embryonic development. The highest levels of expression in the heart coincides with a transition in the expression of isoforms of contractile proteins going from a fetal to a more mature phenotype.

The atrial and ventricular light chain proteins have been used to determine the phenotype of cardiac myocytes derived from an *in vitro* embryonic system. The functional studies performed on ion channels, support the value of this *in vitro* system to study early events in cardiogenesis. Furthermore a system was developed for functional ion channel assessment in concert with immunocytochemistry for phenotyping single isolated cardiomyocytes.

A major part of the thesis was aimed at the identification and characterization of the MLC-2a regulatory DNA sequence. To this end a genomic clone encompassing the complete transcribed part of the gene and 2.0 Kb of the 5' flanking region was isolated and sequenced. In addition 700 bp of the 3' flanking region was analyzed (chapter 6). The structure with regard to the intron-exon boundaries

were compared in various species. Furthermore the homology of various MLC proteins at the amino acid level in mouse, rat and man was determined. The homology of the atrial myosin light chain in mouse and man appeared to be very high (95%). Following identification of the 5' flanking region, this DNA fragment was characterized by transfection studies in rat neonatal cardiomyocytes, skeletal myoblasts and myotubes and non muscle cells. Short proximal promoter fragments of 200 an 450 bp proved to mediate high level, cardiac specific expression of the luciferase reporter gene. The longer promoter fragments (> 450 bp) induced lower luciferase expression levels. The transcription initiation site was identified and appeared to contain a conserved DNA element previously named Initiator. Partial and complete deletion of the Initiator site crippled promoter activity. A map of additional conserved *cis* elements was constructed (chapter 7). The influence of hypertrophic stimuli on MLC-2a promoter activity and ventricular endogenous MLC-2a expression is reported in chapter 8. In cultured ventricular rat neonatal cardiomyocytes promoter activity of the MLC-2a gene was studied in response to different serum conditions, endothelin, phenylephrine and estrogen. A marked induction of luciferase activity was measured upon myocyte treatment with neurohumoral factors. Furthermore the endogenous MLC-2a gene is re-expressed in ventricular myocytes in the absence of serum. The expression levels increase upon stimulation with serum, estrogen, endothelin and phenylephrine. In ventricular tissue of patients suffering from idiopathic dilated cardiomyopathy re-expression of the MLC-2a gene was shown both at the RNA and protein level. Under physiologic conditions the MLC-2a gene is silent in ventricular myocytes. Abnormal MLC-2a expression patterns were found in patients with pump failure and in developing mouse embryos deficient for the retinoid-X-receptor α. The retinoid-X-receptor α knock out experiment indicated the potential role of retinoid acid and its receptors in transcription regulation of the MLC-2a gene in ventricular cardiomyocytes. To analyze the potential role of conserved *cis* elements, that can interact with retinoic acid and its receptors, mutations and deletions were created in an element labeled DR-1 in the MLC-2a promoter. The results indicate a role of the retinoic acid responsive element (DR-1) in the down regulation of the MLC-2a gene in ventricular myocytes. In chapter 10 the major findings of this thesis are summarized and placed in perspective. The consequences of this work for our current understanding of MLC-2 expression in the heart are discussed. The limitations of the methods and the work presented is addressed in addition to what is required for further evaluation of the murine MLC-2a gene. Finally speculations on the future applications of the MLC-2a promoter as a tool transgenic and biotherapy experiments are included.

Samenvatting

Het onderwerp van dit proefschrift is de transcriptie regulatie van de cardiale myosine lichte keten (MLK) eiwitten. In het hart worden een boezem en kamer variant van het eiwit onderscheiden. In hoofdstuk 1 wordt de naamgeving van de myosine lichte keten eiwitten uitgelegd. Voorts wordt de structuur van de sarcomeer besproken evenals de basis principes van de moleculaire biologie. De studies die de basis vormen voor dit proefschrift worden aangehaald om de vraagstelling en studie opzet te toe te lichten. In hoofdstuk 2 worden de verschillende mechanismen van transcriptie regulatie in het hart beschreven. Transcriptie regulatie is een complex samenspel van nucleinezuur moleculen (DNA/RNA) met kerneiwitten (transcriptie factoren). Behalve een bespreking van regulatie mechanismen worden in dit overzichtsartikel een aantal transcriptie factoren, die in het hart actief zijn besproken. Dit gebeurt aan de hand van genen die coderen voor eiwitten, noodzakelijk voor de opbouw van de sarcomeer en waarvan gegevens over transcriptie regulatie zijn gepubliceerd. Strategie en methoden, die gebruikt worden om nieuwe transcriptie factoren op te sporen en te bestuderen worden in een overzichtsartikel besproken (hoofdstuk 3). Door gebruik te maken van technieken uit de moleculaire biologie is het momenteel mogelijk om de interactie tussen DNA of RNA en transcriptie factoren in detail te bestuderen. De exacte base volgorde vereist voor eiwit binding kan bepaald worden en de functionele consequenties van de eiwit-nucleinezuur interactie kunnen gemeten worden. Functionele bepalingen kunnen verricht worden op myocyten in kweek, maar ook in levende organismen.
Door gebruik te maken van de technieken besproken in hoofdstuk 3, hebben we een muizen eiwit opgespoord dat sterk verwant is aan het humane eiwit: SAP49. Het humane eiwit speelt een rol bij het verwijderen van intronen tijdens het produceren van messenger RNA (splice factor). Het muizen gen, genoemd mSAP49, wordt in hartweefsel aangetroffen gedurende de embryonale ontwikkeling, juist in een periode waarin bepaalde foetale eiwitten worden vervangen door meer volwassen isovormen van deze eiwitten (chapter 4).
De ontwikkelingen in de moleculaire cardiologie worden afgeremd door het ontbreken van stabiele cardiomyocyt cellijnen. Een potentiële bron van cardiale myocyten wordt gevormd door een *in vitro* systeem van differentiërende embryonale stamcellen. Deze differentiërende stamcellen vormen klontjes van cellen die spontaan contraheren. Uit deze lichaampjes kunnen cellen geïsoleerd worden die veel overeenkomsten hebben met cardiomyocyten. In hoofdstuk 5 is beschreven hoe we deze cellen hebben gekarakteriseerd aan de hand van het expressie patroon van de MLK eiwitten. Deze vorm van phenotypering werd gekoppeld aan functionele studies naar ion stromen. Deze studie heeft aange-

toond dat immunocytochemie en cellulaire electrofysiologie in één cel gecombineerd kan worden.
Een belangrijk deel van het onderzoek was gericht op de transcriptie regulatie van het atriale MLK gen. Hiertoe werd een genomisch DNA fragment geïsoleerd, waarop het complete MLK-2a gen werd aangetroffen. Dit omvat het complete DNA gebruikt tijdens transcriptie en 2000 baseparen aan de voorkant (5'), en 700 baseparen aan de achterkant (3'). De structuur van het gen werd vergeleken met MLK genen van andere diersoorten, waar het de grootte en de lokalisatie van intron en exon overgangen betreft (hoofdstuk 6). Voorts werd de aminozuur volgorde van MLK eiwitten van muis, rat en mens vergeleken. De homologie tussen muis en mens eiwit is 95%. De promoter van het MLK-2a gen (5') werd geanalyseerd door middel van transfectie studies in hart spiercellen, skelet spiercellen en nier cellen. Korte DNA fragmenten van de MLK-2a promoter (200 en 450 baseparen) gekoppeld aan het "reporter" gen luciferase lieten een hoge activiteit zien. De langere promoter fragmenten lieten een duidelijke afname van de activiteit zien (hoofdstuk 7). De plaats op het gen waar de transcriptie begint werd vastgesteld en bleek samen te vallen met een eerder beschreven regulerend DNA element genoemd: "Initiator". Deletie van dit DNA element maakten de promoter inactief. Andere DNA elementen die eerder herkend werden als bindingsplaatsen voor transcriptie factoren worden besproken. In hoofdstuk 8 worden de resultaten getoond van studies naar het effect van hypertrofe stimuli op de MLK-2a promoter en op re-expressie van MLK-2a in kamermyocyten. In neonatale ventriculaire myocyten in kweek komt MLK-2a tot re-expressie, zelfs zonder serum toevoeging. In de aanwezigheid van serum, oestrogeen, endotheline en fenylefrine wordt een duidelijke toename gevonden in het activiteitsnivo van de MLK-2a promoter. Ook werd re-expressie aangetoond van het MLK-2a gen in kamer weefsel van patiënten met een idiopathische gedilateerde cardiomyopathie. Stimulatie studies met de MLK-2a promoter lieten een toegenomen luciferase activiteit zien na behandeling met endotheline, oestrogeen en fenylefrine. In muizen met een defect retinoid X receptor-α gen, is de embryonale ontwikkeling van het hart gestoord. Op het moleculair nivo bleek kamerweefsel langer dan normaal MLK-2a aan te maken. In hoofdstuk 9 is beschreven hoe met behulp van de MLK-2a promoter dit mechanisme verder kan worden bestudeerd. Door een DNA element dat aan retinoine zuur afhankelijke receptoren kan binden te veranderen of weg te nemen werd aangetoond dat dit element essentieel is voor transcriptie regulatie van het MLK-2a gen in kamerweefsel.
In hoofdstuk 10 is getracht de resultaten van dit proefschrift in perspectief te plaatsen, door aan te geven hoe de resultaten geïnterpreteerd kunnen worden. De beperkingen van de gekozen studie opzet en de noodzaak voor aanvullende studies worden hier uiteen gezet. De mogelijke rol voor de promoter van het MLK-2a gen voor de ontwikkeling van diermodellen voor atriale hartziekten en voor biotherapie wordt hier belicht.

Acknowledgement

The foundation of this thesis is San Diego based. The laboratory of Prof. Ken Chien and his numerous co workers had the patience to guide an ignorant MD around the cliffs and hazards, associated with molecular biology in 1993 and 1994; Ken, I am indebted for the scientific space and confidence you gave me. What puzzled me after each meeting (where I mainly listened) is the level and the number of ideas you can generate in a few minutes. Invaluable help came from Ming Dong Zhou, Steve Kubalak, Sutip Navankassatussas, and Hong Zhu, who suffered on a daily bases; One of a kind Monique Lacorbiere, can show you how to run a zoo, or laboratory, both famous places in San Diego. Kenneth D. Becker thanks for the cover picture and your friendship. Pilar Ruiz Lozano your support, teaching, suggestions, notes and articles got me all the way to this special day. John J. Hunter, your persistence and knowledge were impressive. Denis, the janitor, thanks for throwing all my notes, articles, slides and maps in the trash container, the last day of my stay in San Diego.

During a brief excursion to Rochester, New York my family was impressed by the hospitality and friendship of Prof Robert Kass and his family. Many thanks for the fruitful and pleasant cooperation and good luck with your new position.

Fortunately people like Marc van Bilsen (physiology) help to re-adapt to the Dutch situation and stay focused on doing research the way it should be done and that is step by step. I apologize for the mid summer deadline stress I inflicted on you and your family. The consistency in the way you work and write is remarkable, and hopefully many PhD students can profit from these gifts. As a technician Heleen Erkens helped to get molecular cardiology going in Maastricht. Now Ronald Bronsaer has the position and showed that technicians can work day and night, write articles and give excellent presentations.

The cardiology department in Maastricht, I would like to thank my colleagues (staff and residents) for giving me the opportunity to perform scientific work, in addition to clinical tasks. I realize that the situation where MDs combine the two functions is new, and that there appears to be some injustice in the situation. However I felt fully supported the past three years for which I am grateful. Especially I would like to thank Prof Hein Wellens for training me in cardiology, and for the trust, vision and inspiration to direct me towards molecular cardiology.

Monique Jansens it is all right if you hate me now, but we made it in time after all. Aïsha Knoop, thank you for the art-work.

My special friends from the kindergarten period, Luud, Berry, Joost and Peter, thanks for the fun the past 34 years.
My mother and sisters have been supportive where- and when-ever needed, even in dark and difficult days.
It is doubtful what is most appropriate for my family, gratefulness or apologies, maybe both. Isabella, I am sorry for neglecting you and grateful for letting me do my thing, even if you did not understand every detail, during all the countless hours (and days). Michelle and Pieter, I am sorry for not being there, and thanks for your support.

Publications

1. de Chillou C, **Doevendans PA**, Cheriex E, Rodriguez LM, Metzger J, Pieters F, Smeets JL, Wellens HJ: Echocardiographic wall motion abnormalities and the signal averaged electrocardiogram in the acute phase of a first myocardial infarction. *Eur Heart J* 1993;14:795-798.

2. de Chillou C, Rodriguez LM, **Doevendans PA**, Loutsidis K, van den Dool A, Metzger J, Bar FW, Smeets JL, Wellens HJ: Effects on the signal-averaged electrocardiogram of opening the coronary artery by thrombolytic therapy or percutaneous transluminal coronary angioplasty during acute myocardial infarction. *Am J Cardiol* 1993;71:805-809.

3. de Chillou C, Rodriguez LM, **Doevendans PA**, Loutsidis K, van den Dool A, Metzger J, Bar FW, Smeets JL, Wellens HJ: Factors influencing changes in the signal-averaged electrocardiogram within the first year after a first myocardial infarction. *Am Heart J* 1994;128:263-270.

4. Dekkers GH, **Doevendans PA**, Bessems PJ, Engels LG: Sweet's syndrome: an explosive illness. *Neth J Med* 1994;45:104-106.

5. **Doevendans PA**, Gorgels AP, van der Zee R, Partouns J, Bar FW, Wellens HJ: Electrocardiographic diagnosis of reperfusion during thrombolytic therapy in acute myocardial infarction. *Am J Cardiol* 1995;75:1206-1210.

6. **Doevendans PA**, Cheriex E: Een patiente met collaps, waar rook is...!. *Hart Bulletin* 1995;26:202/204-206.

7. **Doevendans PA**, Hunter JJ, Lembo G, Wollert KC, Chien KR: Strategies for studying cardiovascular diseases in transgenic mice and gene-targeted mice, in Monastersky GM, Robl JM (eds): *Strategies in transgenic animal science*. Washington, American Society For Microbiology, 1995, pp 107-144.

8. An RH, Davies MP, **Doevendans PA**, Kubalak SW, Bangalore R, Chien KR, Kass RS: Developmental changes in beta-adrenergic modulation of L-type Ca2+ channels in embryonic mouse heart. *Circ Res* 1996;78:371-378.

9. Davies MP, An RH, **Doevendans PA**, Kubalak S, Chien KR, Kass RS: Developmental changes in ionic channel activity in the embryonic murine heart. *Circ Res* 1996;78:15-25.

10. **Doevendans PA**, van Bilsen M: Transcription factors and the cardiac gene programme. *Int J Biochem & Cell Biol* 1996;28:387-403.

11. **Doevendans PA**, Becker DA, An R, Kass RS: The utility of fluorescent *in vivo* reporter genes in molecular cardiology. *Biochem Biophys Res Comm* 1996;222:352-358.

12. **Doevendans PA**, Cheriex E, van der Zee R, van den Dool A, van Doorn DJ, de Swart H, Wellens HJ: Risk stratification in the thrombolytic era: results of a prospective study. *Cardiologie* 1996;3:319-323.

13. **Doevendans PA**, Wellens HJ: Atrial dissection for atrial fibrillation: when, how much and where. *J Am Coll Cardiol* 1996;991-993.

14. Kubalak S, **Doevendans PA**, Rockman H, Hunter JJ, Tanaka N, Ross Jj, Chien KR: Molecular analysis of cardiac muscle diseases based on mouse genetics, in Adolph KW (ed): *Human molecular genetics*. Orlando, Academic Press, 1996, pp 470-487.

15. Peters FP, **Doevendans PA**, Erdkamp FL, Van der Ent FWC, De Heer F: Low molecular weight heparin-induced thrombocytopenia and thrombosis. *Eur J Haematol* 1996;56:329-330.

16. Prins A, Tibben A, **Doevendans PA**: Presymptomatisch genetisch onderzoek bij kinderen met een risico voor familiaire hypertrofische cardiomyopathie? *Hart Bulletin* 1996;27:110/113-115.

17. Wehrens XH, **Doevendans PA**, van Dantzig J, Snoep G, Wellens HJ: Acute pericarditis after cardiopulmonary resuscitation. *Am Heart J* 1996;685-686.

18. **Doevendans PA**: Scientific conference on the molecular biology of the normal, hypertrophied and failing heart. *Cardiologie* 1997;4:29-30.

19. **Doevendans PA**: De pathofysiologie van hypertrofische cardiomyopathie. *Ned Tijdschr Geneeskd* 1997;141(1):59-60.

20. **Doevendans PA**, Bronsaer R, Meijer H, Bär FWHM. Moleculaire cardiologie: wetenschappelijk onderzoek en patiëntenzorg. *Hart Bulletin* 1997;28:81-83.

21. Bronsaer RJP, **Doevendans PA**. Keukenzout in het hart: chloor. *Hart Bulletin* 1997;28:149-152.

22. **Doevendans PA**, van Gilst WH, van der Laarse A, van Bilsen M. Molecular Cardiology Part I: Gene transcription in the cardiovascular system. *Cardiologie* 1997;4:225-233.

23. **Doevendans PA**, van Dantzig JM, Schaap C, Meijer H. Molecular genetics of cardiomyopathies. Proceedings Symposium "Molecular cardiology in clinical perspective", AMC Amsterdam 1997: 33-53.

24. Ruiz-Lozano P, **Doevendans PA**, Brown A, Gruber PJ, Chien KR. Developmental expression of the murine spliceosome-associated protein mSAP49. *Developmental Dynamics* 1997;208:482-490.

25. **Doevendans PA**, Ruiz-Lozano P, van Bilsen M. Hunting down nucleic acid binding factors in the cardiovascular system. *Cardiovascular research* 1997 (In press).

26. Wellens HJJ, **Doevendans PA**, Smeets JLRM, Rodriguez LM, den Dulk K, Timmermans C, Vos MA. Arrhythmia risk: Electrophysiological studies and monophasic action potentials. *Book chapter (In press).*

27. van der Laarse A, Ruwhof C, van Wamel JET, van Gilst WH, **Doevendans PA**, van Bilsen M. Molecular cardiology. Part 2: Molecular aspects of cardiac hypertrophy and heart failure. *Cardiologie* 1997;4:328-333.

28. de Muinck ED, van Gilst W, van der Laarse A, **Doevendans PA**. Molecular Cardiology Part 3: Restenosis. *Cardiologie* 1997 (In press).

Abstracts

1. **Doevendans PA,** Bijlsma J: Traumatic luxation of the hip in children. Annual meeting of the Dutch Orthopedic Society. *Meeting report* 1990.

2. **Doevendans PA,** Gorgels A, Wellens HJ: ST-T segment elevation and increase of chest pain accompanies reperfusion during thrombolytic therapy in acute myocardial infarction. *Eur Heart J* 1991;12:265.

3. **Doevendans PA,** Gorgels A, Wellens HJ: Electrocardiographic signs of reperfusion in acute myocardial infarction. *Circulation* 1991;84:117.

4. **Doevendans PA,** van den Dool A, van de Bijl B, Bar FW, Wellens HJ: Utility of the signal averaged electrocardiogram to identify cicumflex coronary infarctions. *J Am Coll Cardiol* 1994;196-A.

5. **Doevendans PA,** Davies M, Kubalak S, An R, Bangalore R, Rubin, Kass RS, Chien KR: Cellular Electrophysiology and Immunocytochemistry of Single Isolated Murine Embryonic Cardio Myocytes: a new method to study cardiogenesis in vivo and in vitro. *Biophysical Journal* 1995;68:A251.

6. **Doevendans PA,** Nguyen V, Becker D, Kass RS, Chien KR: The green fluorescent protein as an in vivo reporter gene in cardiomyocytes and other mammalian cells. *Eur Heart J* 1995;16:S111.

7. Ruiz-Lozano P, **Doevendans PA,** Brown A, Chien KR: A *sex-lethal* related factor (CNR) is a negative regulator of MLC-2v expression. *Circulation* 1995;92:755.

8. **Doevendans PA,** Ruiz-Lozano P, Bronsaer R, Kubalak S, Zhou MD, Bilsen van M, Chien KR: Identification of *cis* regulatory sequences which mediate atrial specificity of the mlc-2a gene in cultured atrial and ventricular cells. *AHA Scien*tific conference on the Molecular Biology 1996 .

9. An R, Davies M, **Doevendans PA,** Kubalak S, Bangalore R, Chien KR, Kass RS: Developmental changes in β-adrenergic modulation of L-type Ca channels in embryonic mouse heart. *Biophys J* 1996;70:A128.

10. **Doevendans PA,** van de Dool A, Konst L, Cheriex EC, Wellens HJ: Significance of bundle branch block after myocardial infarction in the thrombolytic era. *J Am Coll Cardiol* 1997;29;452-A.

11. **Doevendans PA,** Ruiz-Lozano P, Bronsaer R, Chien KR, Bilsen van M: Initiation of transcription by the atrial myosin light chain 2 promoter requires the presence of the INR *cis* element. *J Am Coll Cardiol* 1997;29:87-A.

12. Lutgens E, de Muinck E, Daemen M, **Doevendans PA,** Wellens HJ. Apo-E3 Leiden transgenic mice develop atherosclerotic lesions after HFC diet. *Vascular Biology* 1997.

13. **Doevendans PA**, Hui An R, Kubalak S, Chien K, Kass RS. Embryoid body derived cardiomyocytes provide an alternative source for atrial and ventricular cardiomyocytes. European Society of Cardiology "Myocardial ischaemia, heart failure, and arrhythmia's". Meeting report.1997.

14. Bronsaer R, **Doevendans PA**, van Bilsen M, Wellens HJ. Reporter genes to study cardiac specific gene expression. *Cardiologie* 1997;4:193

15. **Doevendans PA**, van Bilsen M, Bronsaer RJ, Ruiz Lozano P, Chien KR. Regulatory sequence of the atrial myosin light chain 2 gene required for cardiac specific expression. *Mol Cell Cardiol 1997;29:A158.*

16. Oude Ophuis AJM, Bär FWHM, Vermeer F, Janssen W, Engelen D, van Ommen V, de Muinck E, **Doevendans PA**, Dassen W, Wellens HJJ. Angiographic validation of bedside markers of coronary reperfusion in acute myocardial infarction: can unnecessary emergency intervention be reduced in angioplasty centres? *Eur Heart* 1997;18:276.

17. Lutgens E, de Muinck E, Daemen M, **Doevendans PA**, Havekes L, Wellens HJJ. Diet and time-dependent progression of atherosclerosis extent and plaque complexity in APOE*3Leiden transgenic mice. *Eur Heart* 1997;18:599.

18. Lutgens E, de Muinck E, Daemen M, **Doevendans PA**, Havekes L, Wellens HJJ. Longterm diet dependent atherosclerosis progression in APOE*3Leiden transgenic mice. *Circulation*, November 1997 (In press).

19. Ruiz-Lozano P, Merki EE, **Doevendans PA**, Ming Dong Zhou, Chien KR. RXR homodimer and RXR/PPAR heterodimer pathways negatively regulate an atrial marker, MLC-2a, at the transcriptional level in ventricular muscle cells. *Circulation*, November 1997 (In press).