COMPARATIVE PROFILE OF TRANSCRIPTION FACTOR BINDING ACTIVITIES IN VARIOUS TISSUES OF DEVELOPING CHICK EMBRYO

Dissertation submitted to the
Jawaharlal Nehru University for the degree of

MASTER OF PHILOSOPHY

Vibha Rani

SCHOOL OF LIFE SCIENCES
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI – 110 067

2001

JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI- 110 067

CERTIFICATE

The research work included in this thesis entitled **"Comparative profile of transcription factor binding activities in various tissues of developing chick embryo"** has been carried out by **Ms. Vibha Rani** in the School of Life Sciences, Jawaharlal Nehru University, New Delhi. This work is original and has not been submitted so far, in part or full for any other degree or diploma of any other University.

Vibha Rani
VIBHA RANI

(Candidate)

Dr. Shyamal K. Goswami

(Supervisor)

Prof. R. K. Saxena

(DEAN)

Dedicated to...

...my father

ACKNOWLEDGEMENT

A true teacher is the most revered and loved being on the earth. The reverence is internal and heart felt. A true teacher is one who teaches us that the quality of our lives depend upon the quality of our ability to perceive things. The acknowledge of which is wordily method, an expression of gratitude.

I express my immense sense of sincere thanks and gratitude to Dr, Shyamal K. Goswami, my M.phil thesis supervisor. His versatile guidance, relentless support, infallible advise, constant inspiration, novel ideas and helpful discussion enabled me to carry out this work with persistent optism. I acknowledge him for instilling in me immense patience and confidence that are so essential for carrying out scientific research.

I am thankful to Dean sir for providing me all the facility.

I owed the most dept of gratitude to my husband. My gratitude to Sandeep he did for me, can not be adequately communicated here. He infect worked as catalyst. I thanks him for always being there with me with warmth and encouragement, when I was less sure of myself. I express my heartiest thanks to him for his constructive criticism, constant inspiration, strong support and help in all stages of this work.

My sincere thanks are to all CIF staff members for their help and cooperation during the tenure of my work. In particular I would to like to thanks Alexender sir, Khan sir, Mishra sir, Sharma sir for their help and cooperation and providing me all necessary facilities.

I am also thankful to Saini sir for his technical help in photography with utmost accuracy and perfection.

All the colleagues of my laboratory deserve special thanks. How can I forget the support and help I received from Manveen, Neel, Sindhu, Vani and Pampa at all odd moments.

I am grateful to Jamal da for his help. I am thankful to all my batch mates and seniors.

Thanks a lot to Satpal ji for looking after daily cleanliness and washing of glass wares in the laboratory.

I am thankful to my mother-in-law and my sister-in-law, Pinky di for their love and kind cooperation.

I am thankful to my mother, bhyia, bhabhi, sisters and jijus for giving me moral support.

This seminar is tribute to my father who encouraged me to persue higher study. From the core of my heart, grateful to him. His thoughts that " Every difficulty is surmountable if it is approached with will, wish and wisdom", always encourage me to do something. He guided me for choosing the subject and showed the path. I t was he, who molded me in such a fine manner, taught me the essence of life, made me believe in the internal goodness of mankind.

I would like to acknowledge all the chick embryo who sacrificed their lives for my work. May their souls rest in peace.

ABBREVIATION

BP	Base pair
EDTA	Ethylene glycol bis (β-aminoethyl ether)-N,N,N',N'-tetra acetic acid
fm	femto
g	gram
hr	hour
HEPES	N,2- hydroxyethylpiperazine-n'2-ethanesulphonic acid
Kb	kilobase pair
l	liter
TEMED	N,N,N',N'-teramethylethylenediamine
min	minute
ng	nanogram
m	mili
μ	micro
mM	milimolar
M	molar
PMSF	Phenylmethane sulphonyl fluroid
pmole	picomole
U	unit
Ci	Curie
Vol	Volume

CONTENTS

INTRODUCTION

Control of gene expression is one amongst the major regulatory mechanisms operative in a particular cell type. Average eukaryotic cells contain between 5000 to 50,000 genes distributed along their chromosome (1). RNA Polymerase II (Pol II) initiates the transcription of messenger RNA (mRNA) at promoters located 5' of the coding sequence of each gene. These promoters can be quite scattered through the genome, occurring every 200000 base pairs on average in genetically complex eukaryotes. Therefore, efficient gene expression requires that the Pol II must locate those promoters efficiently and selectively to initiates transcription and elongate through the entire coding region of the gene.

1 Control of Eukaryotic Transcription

1.1. General Transcription Factors

The structure of eukaryotic promoters is generally more complex than the prokaryotic ones (1). Most noticeable difference is the existence of lengthy sequences called upstream sites, often hundreds of base pairs apart from the transcription initiation site, that control the rate of initiation. Transcription initiation is a highly complicated process involving a diversity of sequence specific DNA binding proteins. Biochemical efforts to understand the regulation of gene expression began with the discovery of a mammalian nuclear RNA Polymerase activity by Weiss and Gladstone in 1959 (2). These nuclear localized DNA dependent RNA Polymerases (Pol I, Pol II, and Pol III) were subsequently purified based on their ability to synthesize poly ribonucleotide copy of the calf thymus DNA template. These purified DNA polymerases, however, were incapable of selectivity initiating transcription at promoters in vitro. This indicated that additional factors are involved in selective transcription initiation. These factors are called general transcription factors (GTFs), although it is possible that some of them do not participate in transcription of all genes in vivo. The GTFs that so far have been described in the context of Pol II mediated transcription include TF2A, 2B, 2D, 2E, 2F, 2H, 2I and 2J(3) (where TF stands for transcription factor). The polypeptide components of Pol II and most of these GTFs are well conserved among eukaryotes from yeast to human. The entire set of GTFs involved in Pol II mediated transcription

initiation is composed of about 30 polypeptides. Table 1 summarizes the composition and functional properties of various GTFs.

Studies using DNA foot printing, gel electrophoretic mobility shift assays and kinetic assays have revealed that GTFs and Pol II either can assemble in an ordered pathway on a promoter to form an initiation complex or it can bind to the promoter as a pre-assembled holo-enzyme (4)

1.1.a Ordered Assembly of Pol II and GTFs:

In the most general form, the conventional model for transcription initiation by Pol II is characterized by the following series of event (5).

1. Recognition of the core promoter elements by TF2D,
2. Recognition of the TF2D-promoter complex by TF2B,
3. Recruitment of a TF2F,
4. Binding of TF2E and TF2H to complete the pre-initiation complex,
5. Promoter melting and formation of an "open" initiation complex ,
6. Synthesis of the first phosphodiester bond of the nascent mRNA transcript,
7. Release of Pol II contacts with the promoter (upstream clearance), and
8. Elongation of the RNA transcript.

GTF not included in this scheme is TF2A, which can join the complex at any stage after TF2D binding and stabilizes the initiation complex. In this model, regulation can occur during any of the above steps of initiation-complex assembly (3, 4, 5).

Table 1

Subunit Composition and Properties of Human GTFs (1,6,7)

GTFs	SUBUNIT COMPOSITION	MOL.WT	FUNCTION
TF2A	3	37 KD (A) 19 KD (B) 13 KD (C)	All three are Required for activation - Antirepression - Antirepression
TF2B	1	35 KD	-Recruit Pol II/TF2F -start site selection
TF2D	10	38 KD [TBP] 250 KD [TAF2 250] 150 KD [TAF2 150] 135 KD [TAF2 135] 95 KD [TAF2 95] 80 KD [TAF2 80] 55 KD [TAF2 55] 31 KD [TAF2 31] 28 KD [TAF2 28] 20 KD [TAF2 20]	-bind TATA element -HMG-box, bromodomain serine kinase -binds downstream promoter region -histone H4 similarity -histone H3 similarity -histone H2B similarity
TF2E	2	56 KD 34 KD	-Promoter melting, -recruits, modulates activity of TF2 H
TF2F	2	58 KD [RAP 74] 26 KD [RAP 30]	-stimulates elongation -σ homology, cryptic DNA binding, prevents spurious initiation
TF2H	8	89 KD [ERCC3] 80KD [ERCC2] 62 KD [p62] 44 KD [hSSL1] 40 KD [cdk7] 37KD[CYCLIN- H] 34 KD [p34] 32 KD [MAT-1]	-3'-5' helicase, excision repair -5'-3' helicase, excision repair -excision repair - excision repair -CTD-kinase, component of cdk7-activating kinase -cdk7 partner -, homologous to hSSL1 -CAK assembly family

1.1.b. DNA Sequence Elements Directing Transcription by Pol II

Three distinct classes of DNA sequence elements direct transcription by PolII (1,8). The first family includes the core or basal promoter elements found near the site where Pol II initiates transcription. Two classes of core elements have yet been identified (9). The TATA element (consensus TATAa/tAa/t), located ~ 25-30 base pairs upstream of the transcription start site, and the initiator motif, a pyrimidine rich sequence (consensus PyPyA+NT/ApyPy) encompassing the transcription start site(10). The TATA and the initiator element, which serve to nucleate the initiation complex, are recognized by components of the basal transcription machinery. They can function independently or synergistically and cellular promoters may contain one these elements, both of these elements or neither. Two other families of cis-regulating elements are the promoter proximal element, situated fifty to few hundred base pairs upstream of the start site, and the promoter distal elements (enhancer) found upto tens of thousands base pairs away form transcription start site. Both of these elements contain binding sites specific for proteins that modulate transcription of the gene concerned and are thus identified as gene specific transcription factors, which act either as activator or repressor proteins. Activator (or repressor) protein complexes influence the rate of transcription of an adjacent gene by Pol II. These activator proteins are believed to influence transcription in part by affecting stable initiation complex formation (11). Activator proteins can stimulate transcription directed by the TF2D multi-subunit complex in vitro, but they do not stimulate reactions directed by recombinant TBP alone. These observations led to the concept that the TAFs subunits of the TF2D are the components of GTF that interact with the activators(12,17). Several other factors have been described that are distinct from GTFs but also appear to play an important role in transcriptional regulation. Among these are the five SRB proteins (13,14), which are integral component of Pol II holoenzyme. A genetic screen for mutations that suppress the conditional phenotypes of Pol II carboxy-terminal-domain (CTD) mutants lead the discovery of SRB genes. Srb 8-11 (also known as Ssn 5, 2, 3, 8), additional subunits of the holoenzyme are required for repression by several repressors (15,16). Srb 10 and 11 form a cyclin-dependent kinase (cdk)- cyclin pair that phosphorylates the Pol II carboxy terminal domain (CTD)(18,19).

1.1.c Role of TAFs in Pol II Mediated Transcription

In addition to TBP, TF2D contains at least eight additional subunits, generally termed as the TBP-associated factors [TAFs], ranging in size from 20 to 250KD. The TAFs are a phylogenetically conserved set of proteins identified in humans, *Drosophila* and Yeast(1,2,). The TAFs are required for activator-dependent transcription stimulation in vitro (20). It has been demonstrated that addition of the TAFs to TBP restors transcriptional activation, indicating that one or more of the TAFs might function as coactivators. It has been found that different classes of activation domains present in the gene specific transcription factors bind distinct TAFs. For instance, Gln-rich activator sp1 and *bicoid* bind TAF10. Acidic domain of p53 and VP16 bind TAF40 and TAF60. The estrogen receptor binds hTAF30, and hTAFs55 appear to bind multiple activators including the CAAT transcription factor (Pro-rich) (21). Therefore it appears that TAFs play a major role in communicating between the basal transcription machinery and the upstream activators (4,20).

1.2. Gene Specific Transcription:

Modulation of gene expression underlies the basic cellular processes such as cell differentiation, organ development, stress response etc. Although eukaryotic gene expression is initiated by the transcription apparatus assembled at the core promoter region (TATA sequence/ Down stream promoter elements), the extent of transcription is controlled by the gene specific transcription factors operating at the upstream regulatory regions of the promoter (3,22). The gene specific transcription factors constitute the largest and the most divergent class of DNA-binding proteins. Gene specific transcription factors mediate their effects by diverse mechanisms such as tissue restricted activity, existence of different iso-forms, formation of functional complex by hetero and homo dimerization etc. They are characterized by their specific binding to distinct DNA elements found in the regulatory regions of different genes. Over the last decade a number of such transcription factor families have been characterized.

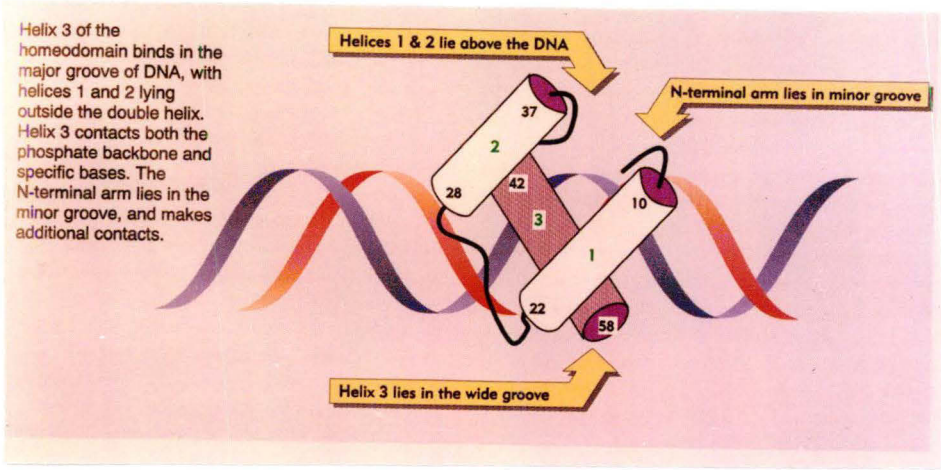
1.2.a Helix-Loop-Helix Family of Transcription Factors

The helix-loop-helix family of transcription factors are characterized by a common structural motifs i.e. a stretch of 40-50 amino acid residues forming two amphipathic α-helices separated by a linker region (the loop) of varying length. The members of this family form both homodimer and heterodimers by means of interactions between the hydrophobic residues on the corresponding faces of the two helices (1,23). The helical regions are 15-16 amino acid long and each contains several conserved residues. The loop is probably important only for allowing the freedom for the two helical regions to interact independently of one another.

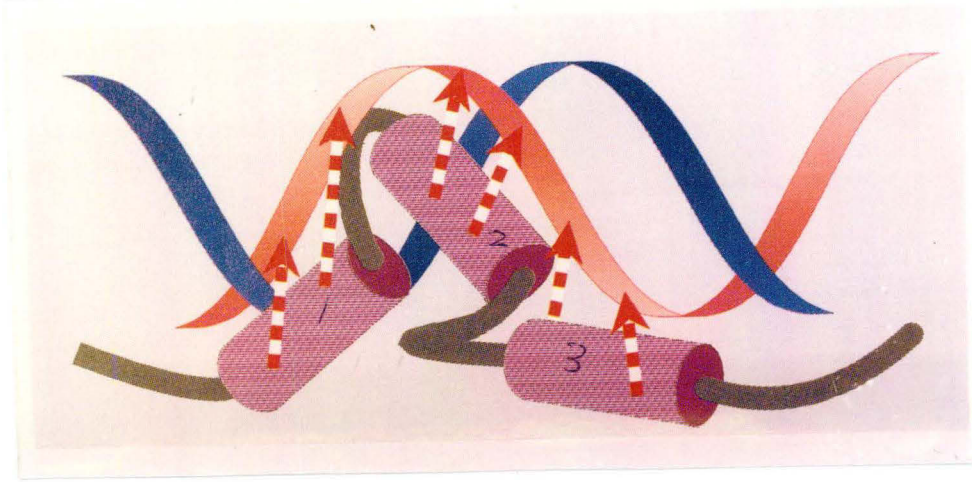
Most HLH proteins contain a region adjacent to the HLH motif itself that is highly basic and is required for its DNA binding activity. Members of the group with such a region are called bHLH (basic helix-loop-helix) proteins. Two distinct classes of bHLH family of transcription factors have extensively been characterized. One family includes the two ubiquitously expressed proteins E12 and E47 that bind to the E-box (CANNTG) element originally identified in immunoglobulin gene enhancer . The other is the myogenic transcription factors viz. myoD, myogenin, myf5 and MRF4 that also bind to the same (E-box sequence) found in the regulatory regions of various muscle specific genes (24). The myogenic bHLH proteins are expressed only in muscle cells and they bind to their target sequence as a hetero-dimer with the E12/E47 proteins. A number of other bHLH transcription factors have also been extensively been charaterized in the context of cell growth/tumour progression (proto-oncogene c-*myc)* and embryonic development (*twist, daughter*less in *Drosophila*).

1.2.b Homeodomain Family of Transcription Factors

The homeodomain is a DNA-binding motif that is present in a large family of eukaryotic transcription factors. The homeobox is a conserved sequence coding for a domain of 60 amino acids originally identified in Drosophila homeotic genes that determine the identity of body structures (59). Subsequently, similar motifs were found in a large number of genes in other eukaryotes. It is now believed that these genes are widely involved in embryonic development and are well conserved in evolution (60).



Homeo domain DNA binding protein



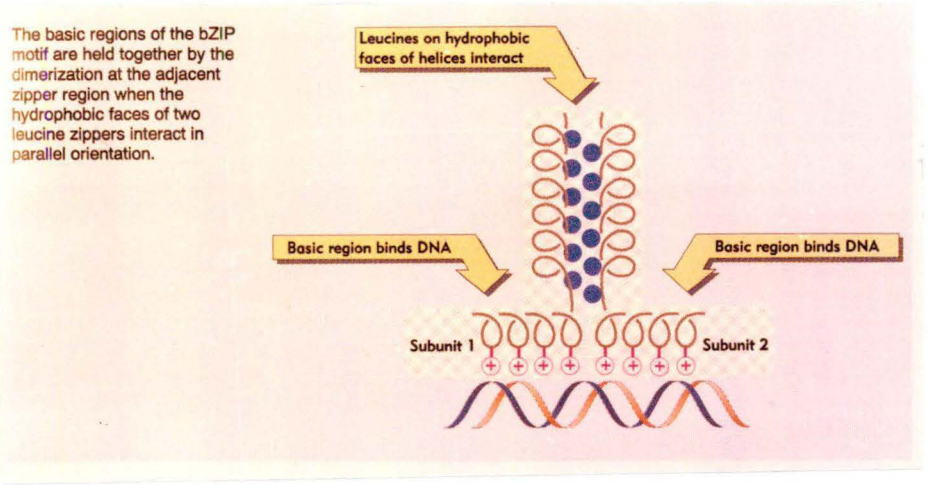
Zinc finger DNA binding protein: Zinc fingers form α-helices that insert into the major groove, associated with β-sheets on the other side. Three fingers are indicated in figure.

1.2.c Helix-Turn-Helix Family of Transcription Factors

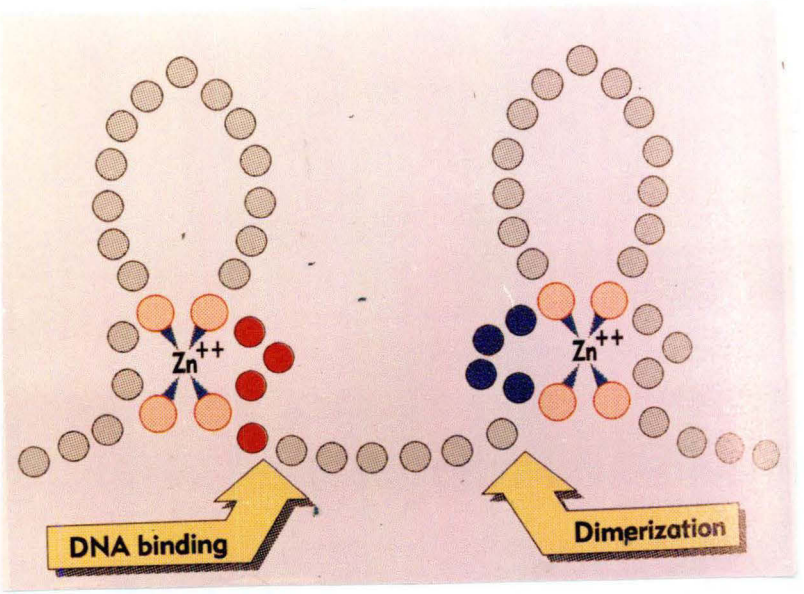
The helix-turn-helix (HTH) structure is the first DNA recognition motif discovered and it has now been determined to be present in a large number of DNA binding proteins in both prokaryotes and eukaryotes (25). The most well characterized member of the family are λ-cro protein, the Escherichia coli CAP protein and the λ repressor. Originally, comparison of these three proteins revealed this conserved recognition motif consisting of α-helix, a turn and a second α-helix. Structures are now available for several other HTH proteins such as the Lac repressor, 434 repressor, 434 cro, Trp repressor and Fis protein (25). The DNA binding domain of LexA also contains a related two helix motif, but the turn region is more extended than that in the other proteins. The HTH motif includes a number of conserved residues involved in stabilizing the HTH structure. The HTH proteins also contain additional domains that have important role(s) in their regulating activity such as dimer formation and binding allosteric effectors.

1.2.d. Zinc Finger Family of Transcription Factors

Zinc finger motif, first discovered in Xenopus transcription factor 3A (TF3A), is also a major structural motif involved in protein-DNA interaction. Zinc finger proteins are involved in many aspects of eukaryotic gene expression such as cell differentiation, transduction of growth signals, development and organogenesis. The zinc finger motif is structure in which a small group of conserved amino acids bind a zinc ion and form a finger like structure (26). They usually contain tandem repeats of the zinc finger motifs, with each motif containing the sequence pattern Cys-$X_{2or\ 4}$-Cys-X_{12}-His-X_{3-5}-His. The motif is identified from the loop of amino acids that protudes from the zinc-binding site and is described as the Cys2/His2 finger. The finger itself comprises of ~ 23 amino acids and the linkers between fingers is usually 7-8 amino acids. Two Cysteines, which are near the turn in the β-sheet region, and two Histidines, which are in the α-helices, coordinate a central zinc ion and hold three secondary structures together to form a compact globular domain (2). The crystal structure shows that the zinc fingers bind in the major groove of B-DNA and wrap partway around the double helix. Each finger has a similar way of docking against the DNA and makes base contacts with a three-base-



Leucine zipper DNA binding protein



Steroid receptor: The first finger of a steroid receptor controls specificity of DNA-binding (red dots in the figure): the second finger controls specificity of dimerization (blue dots in the figure).

pair sub-site. Besides the other classical zinc finger proteins members of the steroid receptor superfamily belongs to this class (27). They are transcription regulators activated upon binding to various steroids and similar compounds. In contrast with the tandem repetition of the Cys2/His2 type of the zinc fingers steroid receptors often have non-repetitive fingers. Genetic and biochemical studies done in a number of different laboratories revealed that these proteins contain separate domains for hormone binding, DNA binding and for transcriptional activation.The glucocorticoids (GR) and estrogen receptors (ER) are among the members of the family that has extensively been characterized. Both GR and ER have two finger motifs, each with a zinc atom at the centre of the tetrahedron of Cysteins. Two glucocorticoid receptors dimerize upon binding to DNA and each engages a successive turn of the major groove. This fits with the palindromic nature of the response element.

1.2.e Leucine Zipper Family of Transcription Factors

The Leucine Zipper (LZ) family of transcription factors has major role in cell growth and differentiation. LZ motif was first discovered as a conserved sequence pattern in CREB and AP-1 transcription factors and it is now clear that this motif appears in a wide variety of transcription factors in fungi, plants and animals (26). Leucine zipper sequences are characterized by a heptad of Leucine residues over a region of 30-40 amino acids. The LZ domain is adjoining to a basic domain that binds to the DNA. The LZ region mediates homo (self) or hetero-dimerization (with other members of the family) and the basic region contacts the DNA. The two Leucine Zipper in effect form a Y-shaped structure in which the zipper comprises the stem and the two basic regions bifurcate symmetrically to form the arms that bind DNA.

1.2.f β-Sheet Motifs

Moajority of the DNA binding proteins that have been structurally characterized bind with an α-helix in the major groove. Nonetheless, the Met J, Arc and Mnt repressors belong to a family of prokaryotic regulatory proteins that uses an anti parallel β-sheet for DNA binding. Met J binds as a tetramer to an 18-base-pairs DNA site. In this complex a dimer binds to each half site. E.coli IHF protein is supposed to use β-sheets for site specific recognition (4,26).

1.3 Transcriptional Regulation of Cell Differentiation and Organogenesis

Developing embryo has long been of interest for studying cell differentiation and organogenesis. In higher organisms, in addition to a common pool of genes that are expressed in all the organs, a distinct set of genes is expressed in a tissue-restricted manner (28). This differential profile of gene expression is primarily achieved at the level of transcription and thus has been investigated as a paradigm for studying molecular basis of cell differentiation and organ development. Those studies have revealed that the tissue specific gene expression in a particular organ depends on its developmental history and is primarily achieved by the combinatorial effect of a number of tissue specific and ubiquitous transcription factors (29). Among the various organs that have been investigated for his purpose, studies on the molecular basis skeletal/cardiac muscle development have significantly contributed to our present understanding of the transcription control of cell differentiation.

1.3.a Skeletal Muscle Development

During the last decade a significant progress has been made in understanding the molecular basis of skeletal muscle development and as consequence, skeletal myogenesis has become a paradigm for understanding the molecular basis of the establishment of cell lineage and cell differentiation (30). Specification and differentiation of skeletal muscle cells in various species is critically dependent on a family muscle-specific helix-loop-helix (bHLH) transcription factors, frequently referred to as the muscle determination factors (MDFs)(31). While in *Drosophila* there is one such factor known as *nautilus*, in mammals, as many as four such factors (MyoD, Myf-5, myogenin, and MRF4)(32) has yet been identified (33). In developing embryo, skeletal muscle originates from the somites, which is formed by the segmentation of the paraxial mesoderm (34,37). During mouse embryogenesis, the myogenic bHLH factors show overlapping, but distinct expression patterns in the skeletal muscle lineage (35,36). Myf5 is the first member of the family to be expressed, myogenin is expressed, about a half-day later MRF4 / MyoD are expressed beginning at Day 9.5 and 10.5 p.c

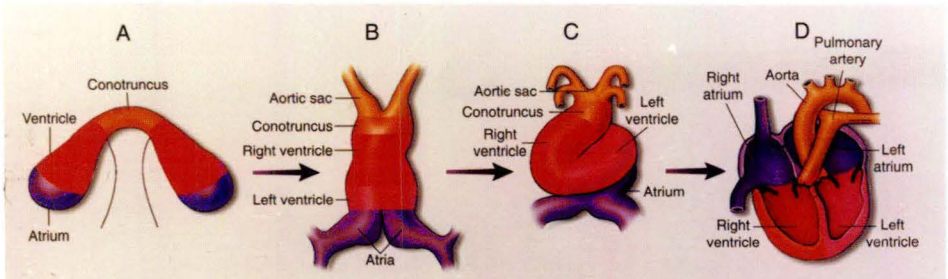
respectively. The role of these factors in muscle development have been confirmed by gene knock-out studies, which have shown that MyoD and Myf5 play redundant roles in the generation of myoblast, whereas myogenin mediates their terminal differentiation in to myotubes (38). MRF4 is expressed late in muscle differentiation and may share function with Myogenin. The myogenic bHLH proteins activate muscle gene expression by binding to the consensus sequence CANNTG (N, any nucleotide) as heterodimer with ubiquitous bHLH proteins known as E-proteins (39). This DNA sequence, commonly referred to as E-box, is found in the control regions of many but not all-skeletal muscle genes. While E boxes are required for transcriptional activation of many skeletal muscle genes, they are not by themselves sufficient and are dependent on the adjacent binding of other factors to activate skeletal muscle gene transcription. Myocyte Enhancer Factor (MEF2) binding sites (TTATTTTTA) often occur in conjunction with the E-boxes in the regulatory regions of various muscle genes and have been shown to be required for transcriptional activation of those genes in skeletal muscle cells (40,41). In addition to the MRDs a number of other transcription factors such as the homeodomain transcription factor Pax3, LIM domain transcription factor MLP (muscle LIM protein), MADS box transcription factor MEF2 also contribute to the skeletal myogenesis (42). Pax3 is a member Paired-box gene family that encodes transcription factors containing paired-domain. Splotch mice carrying Pax3 mutation lack limb muscles as a consequence of impaired migration of muscle progenitors from the lateral dermomyotome, although they are capable of differentiation (33). Splotch-Myf5 double mutant mice fail to express MyoD in the myotome and lack all body muscle but not head muscles. Thus Pax3 is an essential upstream regulator for MyoD expression in precursor of the body muscle but not of the head muscle (40,49). A novel protein, called muscle LIM protein (MLP), has also been characterized in the context of muscle differentiation. The MLP gene is expressed in skeletal muscle and its expression is significantly up regulated during myoblast differentiation where it accumulates in the nucleus of differentiating myotubes (43).

In addition to the E-box sequence, the regulatory regions of a number of muscle specific genes also contain the target site for the Myocyte Enhancer Factor-2 (MEF2) (44). While *Drosophila* has one MEF-2 gene, vertebrates have four iso-genes (MEF2A-

2D), which generate proteins containing the highly conserved MADS and the adjacent MEF2 domains. MEF2 proteins bind as either homo- or heterodimers to an A-T rich DNA consensus sequence (TTATTTTTA) present in the control region of most muscle specific genes. It is believed that MEF2 transcription factors not only activate target genes by binding to their own DNA recognition site but also by protein-protein interactions with myogenic bHLH/E12 heterodimers which are bound to the E-box consensus sequence (45). This interaction of MEF2 factors and bHLH heterodimers occurs through their respective DNA-binding and heterodimerization motifs. In this complex MEF2 probably act as an essential cofactors for muscle-specific gene activation (46,47).

1.3.b Transcriptional Control of Cardiac Development

Since a functional circulatory system is essential for the progress of embryogenesis, heart is the first organ to develop and cardiac myocytes are amongst the earliest differentiated cell types (48). In vertebrates heart is initially formed as a tubular structure that subsequently undergoes extensive morphogenesis leading to the formation of the four- chamber organ (49). During embryogenesis, heart derives from the anterior portion of the lateral plate mesoderm (50). Induced by signals from the underlying endoderm, splanchnic mesodermal cells ventral to the pericardial coelom become specified to the cardiac fate and differentiate in to bilateral precardiac mesoderm (51,58). In the mouse, by 8.0 days post coitus (d.p.c) the embryo folds and the bilateral heart primordia migrate to the ventral midline and fuse with each other to form a single heart tube. The straight tube then undergoes looping and septation. At around 9.5 d.p.c. the atrial portion shifts dorsally and to the left and boundaries between the common atrium, the primitive ventricle (the future left ventricle) and the bulbas cordius (the future right ventricle) become prominent, principally due to the formation of endocardial cushion. Septation in the common atrium begins around 10.5 d.p.c. All of cardiac valves form by 13.0d.p.c. and the embryonic heart acquires a definitive four chamber structure by 14.0 d.p.c. In the developing chick embryo, the cardiogenic fate is assigned to the mesodermal precursors as stage 4 (20 hours) and at about stage 10 (33 hours) the tubular heart is formed (55,61). Tubular heart is capable of rhythmic



Schematic of the major morphologic events in vertebrate heart formation. Cells from the precardiogenic mesoderm are specified at the embryonic primitive streak stage (A) to form the various structures of the linear heart tube (B). Later, the heart loops rightward (C) and finally septates and forms the various regions and chambers of the looped and mature heart (D).

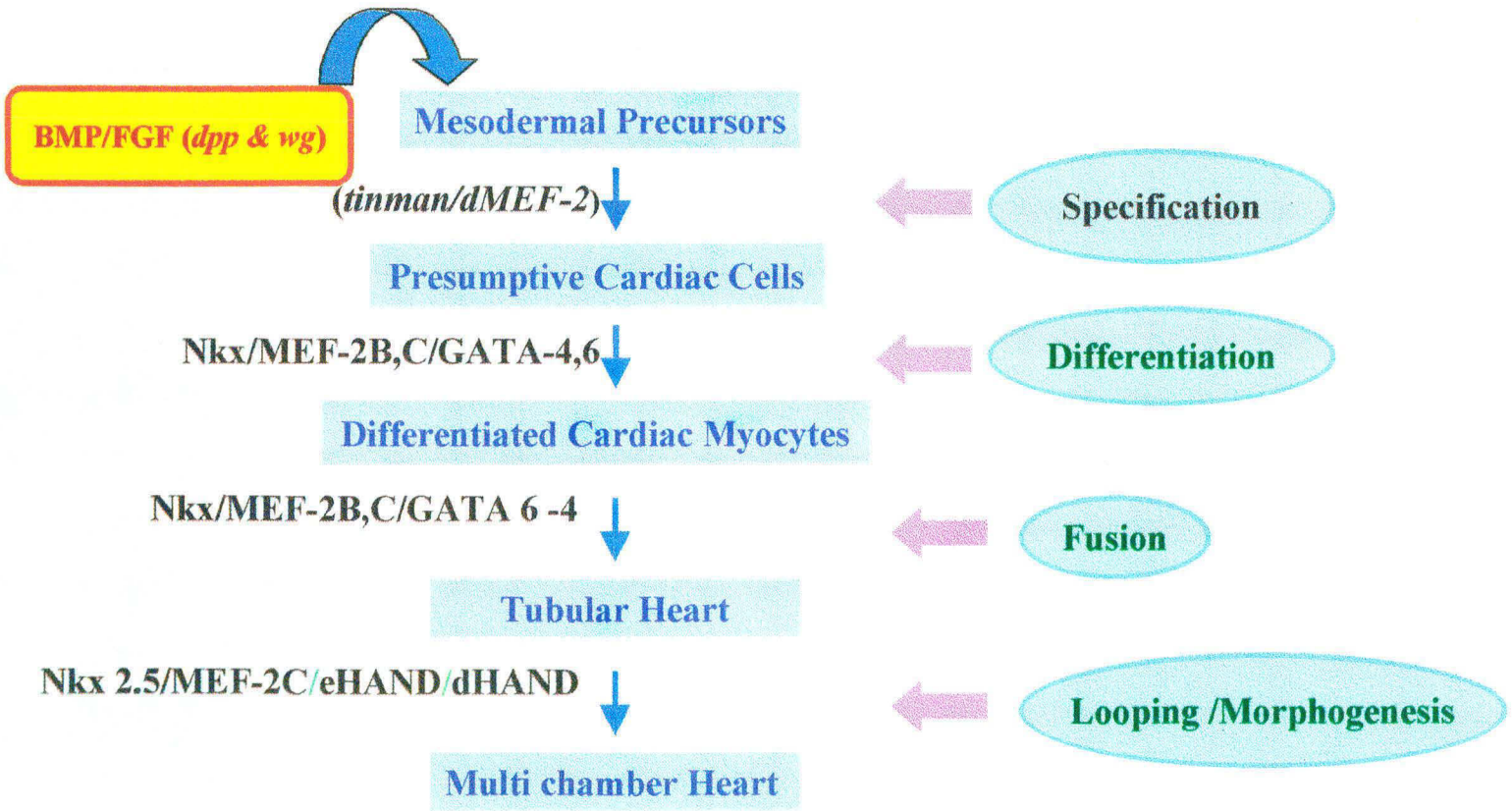
Adopted from Heart Development: edited by Richard P. Harvey & Nadia Rosenthal

contraction and contains at least three distinct cell types viz. endocardial- endothelia, ventricular myocyte and atrial myocytes (56). *Drosophila* heart, is a simple tubular structure consisting of two types of cells and is known as dorsal vessel (57). Since vertebrate heart is also formed as a tubular structure in the beginning, dorsal vessel in *Drosophila* can be perceived as a primitive form of vertebrate heart. Furthermore, both Drosophila and the vertebrate hearts originate from similar region of the mesoderm. Molecular genetics of Drosophila development has thus significantly contributed to the general understanding of the molecular basis of heart muscle development (58).

In spite of the biochemical and structural similarities between the cardiac and skeletal muscle cells, the basic difference between these two muscle type is the absence of MyoD family of master transcription regulator in skeletal muscle cells (59,60,61). Based on the genetic analysis of the Drosophila and zebrafish development as well as analysis of the regulatory regions of various muscle specific promoters, a number of transcription factor such as Nkx2.5, GATA-4/5/6, MEF2 and e/dHAND have been implicated to cardiac muscle development (51,62). Discovery of *tinman* gene in *Drosophila* that is essential for mesoderm induction as well as cardiac and visceral muscle development, led to the isolation of its vertebrate homologues viz Nkx2.3, Nkx2.5, Nkx2.6, Nkx2.7, Nkx2.8 and Nkx2.9 from a number of species such as mouse, chicken, zebrafish (63) and frogs(64). Among these Nkx family members, only Nkx2.5 is consistently expressed in all species and the cardiac expression of others is either restricted or they are expressed only at a certain stage of development (65,77). The homeodomain protein Nkx2.5/Csx is expressed in precardiac mesoderm and in atrial and ventricular myocyte throughout the development. *Tinman*, Nkx2.5 and other members of this family act as a transcription factors that bind to the DNA element AAGTG as monomers. Targeted interruption of Nkx2.5 function in transgenic mouse embryos does not prevent the formation of the primitive heart tube but lacks looping morphogenesis, indicating that it is not necessary for myocyte commitment. It is hypothesized that Nkx family harbors a common "Nkx code" and it is their unique combinatorial expression that directs various cell fates (66,75).

A second set of nuclear factor important in vertebrate heart development is members of the MEF2 family (78). Differentiated cardiac and skeletal muscle cells express a battery

DEVELOPMENT OF HEART MUSCLE



of genes essential for its contractile function. Among various cis-regulatory elements frequently found in the regulatory region of muscle specific genes is C/TTA(A/T)$_4$, target site for the MEF2 (Myocyte Enhancer Factor-2) family of transcription factors. Two members of this family, MEF2C and MEF2B are expressed in mouse cardiac primordia and throughout heart development, suggesting that they play an important role in cardiac myocyte differentiation. Targeted disruption of MEF2C results in normal growth until E9.0 following which it shows defects in initiation in rightward looping and right ventricle formation. A third multigene family involved in cardiac commitment and differentiation is the GATA family of zinc finger proteins. Member of GATA family of transcription factors bind to the DNA element, for which they are named, A/T**GATA**A/G found in the regulatory regions of various genes expressing in different tissues (67). GATA factors are generally implicated in cell lineage determination in vertebrates (52,67). Six members of GATA family (GATA1-6) have been identified. GATA1/2/3 are involved in hematopoisis and brain development while GATA4/5/6 are expressed in heart, gut epithelium and lung (68). During mouse embryogenesis, GATA 4/6 mRNAs can be detected in the precardiac mesoderm and visceral endoderm as early as E7.5 and it remain afterwards. However GATA5 mRNA is detectable in the atrial and ventricular chambers at E9.5, it becomes restricted to the atrial endocardium at E12.5 and subsequently it is silenced. Embryos homozygous null for GATA4 form cardiac myocyte expressing contractile protein genes but fail to migrate them to form heart tube (69,72). In addition, these embryos have defects in cardiac and ventral morphogenesis. The transcription factor GATA6 is first expressed at the blastocyst stage in part of the inner cell mass and in the trophectoderm. The second wave of expression is in parietal endoderm, mesoderm and endoderm that form the heart and gut. Inactivation of GATA6 leads to lethality shortly after implantation (70,72,73). Recently several basic region helix-loop-helix (bHLH) domain proteins have been identified in the developing myocardium. Two related bHLH protein, eHAND and dHAND are expressed in the bilateral cardiac primordia. When antisense oligo-nucleotide either of these transcripts are used to block the appearance of the protein during chick embryonic heart development, there is no abnormal phenotype. When anti-sense oligos to both are used, however heart development is arrested at the looping heart

tube stage. dHAND is expressed in all cardiac myocytes whereas eHANDs shows a transient pattern of expression predominantly in one ventricle; suggesting it may be involved in segment-specific gene expression during development (71,74,75,76).

1.3.c Transcriptional Control of Lung Development

Fetal lung development is a highly complex and orchestrated process involving growth, differentiation, branching morphogenesis and synthesis of the lipoprotein surfactan (77). Members of the fibroblast growth factor (FGF) (78,79,), transforming growth factor -β (TGF-β), and epithelial growth factors (EGF) families, as well as the secreted morphogen sonic hedgehog (shh), appear to play important roles in lung branching morphogenesis and epithelial cell differentiation (81,82,83). The Gli family of zinc finger containing transcription factor, viz. Gli1, Gli2 and Gli3 are expressed in splanchnic and lung mesenchyme in different spatiotemporal patterns during embryonic lung development. In mice, homozygous for a targeted deletion of the Gli3 gene, the lungs are reduced in size and there is a disproportional decrease in length of the left lobe (84). By contrast, in Gli2-1 mice, both lungs are markedly reduced in size because of defects in lung proliferation (85). In Gli2-/-Gli3+/- mice, there is failure of separation of the trachea and oesophagus; the lung is more hypoplastic and does not separate into right and left lobes. On the other hand, mice homozygous for deletion of both Gli2 and Glui3 genes mostly die at E10.5 and those survive until E13.5-E14.5 lack lungs, trachea and oesophagus (86). These finding indicate the essential role of mesenchymally expressed Gli2 and Gli3 in development of foregut endoderm into trachea, oesophagus and lungs The finding that the lung and tracheal /oesophagal phenotype of the Gli2 -/- and Gli2-/-/Gli3+/- embryos is similar to that of Shh -/- mice suggests that transcription factors Gli2 and Gli3 mediate the effects of Shh on foregut development into these structures (87).

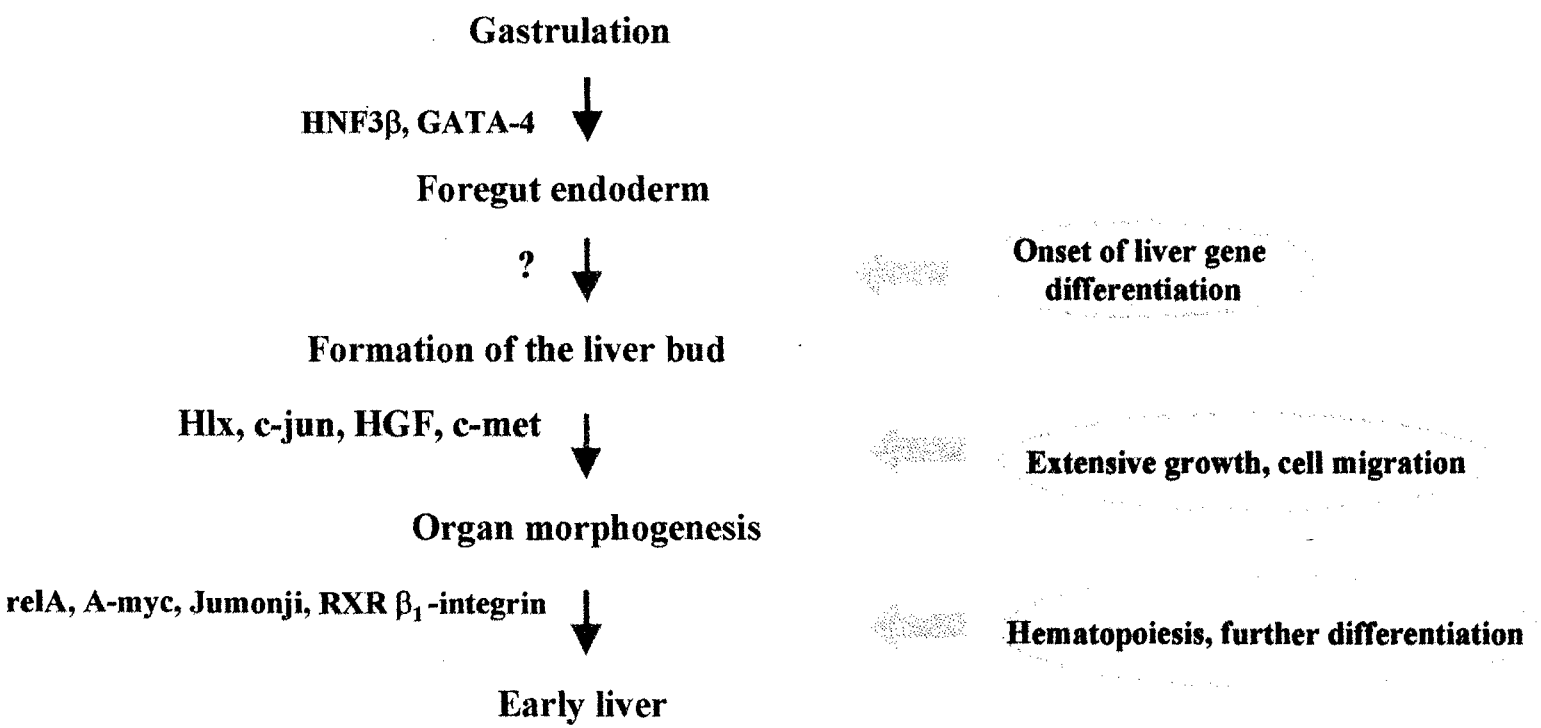
In addition to Gil family of transcription factors, several other transcriptional regulators have also been implicated in lung development and morphogenesis (88). Retinoids appear to have a role in maintaining the differentiated state of the respiratory epithelium (89). These effects are presumably mediated by retinoid receptors, which are members of the nuclear receptor superfamily. There are a number of three related retinoic acid

receptors existing in several isoforms. RAR is expressed ubiquitously throughout the developing mouse lung (90). In a detailed analysis of the effects of targeted deletions of various members of the RAR and RXR subfamilies, it was found that, whereas RXR --/- mouse embryos have defects in heart and liver development, there is no obvious lung phenotype (91). However, mice carrying targeted deletions of both RAR- and RAR genes manifest failure of separation of oesophagus and trachea, agenesis and hypoplasia of the left lung, hypoplasia of the right lung and defects in lobation (92). N-Myc, a nuclear proto-oncogene of the basic helix-loop helix family, has also been implicated in lung development. N-Myc -/- embryos die at E11.5 and have major defects in heart, central nervous system, lung and gut development (93,94). Thyroid transcription factor-1 has also been implicated in lung development (95,96) The glucocorticoid receptor (GR) is a member of the nuclear receptor superfamily of transcription factors which upon binding to glucocorticoid dissociates from a complex with heat shock proteins, enters the nucleus and binds to the glucocorticoid response elements (GREs) in the regulatory regions of target genes (95,98). Expression of the GR mRNA is detectable in embryonic lung and appears to be localized mainly to mesenchyme (99,100). Mice homozygous for targeted deletion of the GR gene (GR-/-) die within several hour of birth as a result of respiratory failure emphasizing the importance of GR in lung development (101,102). A role of cyclic AMP Response Element (CRE) Binding Protein (CREBP) in fetal lung development is suggested by the finding that mice homozygous for targeted deletion of the gene die within 15 min of birth and the lung of these animals contain low levels of the surfactant proteins SP-A , SP-B and SP-C (103,104,).

1.3.d Transcriptional Control of Liver Development

Hepatocytes undergo distinct phases of differentiation as they arise from the gut endoderm, coalesce to form the liver and mature by birth. Hepatic differentiation begins in the ventral foregut endoderm immediately after the endodermal epithelium interacts with the cardiac mesoderm, in the ~ 6 somite embryo (105,106,107). This hepatic induction step occurs at 8-8.5 days of gestation in the mouse (108,109). Gene in-activation studies in mice have shown that the transcription factors hepatocyte nuclear factor3β (HNF3β) and GATA4 are required for the morphogenesis of the foregut

LIVER DEVELOPMENT



endoderm (110), the precursor of the liver. HNF3β contains the novel 'winged helix' DNA-binding motifs and GATA4 contain $(cys\text{-}cys)_2$ zinc fingers (111). Various tissues other than the liver, such as the lung and pancreas are derived from the fore gut endoderm (112), and these tissue express HNF3 and GATA family members (113,114). It appears that HNF3 and GATA (115) proteins initially open up and remodel chromatin sites in the endoderm, thereby priming target genes for activity during tissue specification (116,). Homozygous mutations in a number of other genes such as c-jun (117), Hlx (118), scatter factor/hepatocyte growth factor (HGF) (119) and c-met give embryonic lethal phenotype with regards to mouse liver development. In each, the hepatic endoderm is specified in terms of expressing genes such as albumin, but the fetal liver is very small, hepatocytes exhibit hypoplasia and sparsely fill the liver matrix.

Materials and Methods

2.1 Polymerase Chain Reaction:

Enzymes	Taq DNA Polymerase
10X Buffer	750mM Tris-HCl (pH 8.8 AT 25°C), 200 mM $(NH4)_2SO_4$,
Primers Used	T3/T7, P1/P2, M1/M2

for 100 µl reaction-

10 X Buffer	10µl
dNTP (genetix 10mM)	8µl
Primer Mix (50 ng/ul)	4µl (M1/M2) and 3µl (PI/P2 or T3/T7)
25mM MgCl2	1 mM with MI/M2 and 3mM with T3/T7 or P1/P2
Template	1µL (5pg)
Enzyme	1µL

Add H_2O to make up the volume 100µl

Mix well

Made 25ul aliquots. Mineral oil is added.

The cycle was as follows

1. 95 °c ..... 5 min
2. 95°c ..... 1 min
3. 55°c ..... 1 min
4. 72°c ..... 30 sec
5. 34 times to 2
6. 72°c ..... 10 min
7. End

2.2 Purification And Precipitation Of DNA

After PCR the DNA is precipitated as follows

1. Added one volume of TE- saturated phenol-chloroform.
2. Vortexed for one min and centrifuged it at 12,000 rpm for 5 min.
3. Transferre the upper, aqueous phase to a fresh tube,
4. Added one volume of chlcroform and isoamyl alcohol (Qualigen) [24:1]

5. Vortexed for one min and centrifuged at 12,000 rpm for 5 min.
6. Removed the upper aqueous phase to another tube.
7. Volume is then measured, and added 1/10 volume of 1M NaCl and 2.5 volume of ethanol (Merck)
8. Mixed well, and kept the tubes in -20° for 2 hr.
9. Again the tubes was centrifuged at 14,000 rpm for 10 min.
10. Supernatant is discarded.
11. The pellet is washed with prechilled 70% alcohol (Qualigen), dried
12. Resuspended the pellet in TE (Tris-HCl, pH 8.0 10mM, EDTA, pH 8.0 1mM).

2.3 Isolation Of Plasmid DNA By Miniprep Method:-

The plasmid mini-prep method allows for the rapid isolation of small amounts of plasmid DNA.

1. Inoculated 1ml of LB media (Hi media) containing 50ug/ml ampicillin (Biovac injections, stock solution concentration is 100mg/ml) with a single bacterial colony.
2. Placed 1ml of overnight culture in to a microcentrifuge tube, and centrifuged the tube at 5000 rpm for 10 min.
3. Removed the supernatant by pouring so that about 50ul is left.
4. Resuspended the pellet by vortexing.
5. Added 300ul of mini-prep buffer (0.1N NaOH (BDH), 0.5% SDS(sigma) and TE) Mixed it by vortexing the tube and kept on ice for 5 min.
6. Added 150 µl of potassium acetate (Qualigen) solution, pH 4.8 and kept on ice 5 min.
7. Vortex it and centrifuged at max speed for 5 min.
8. Transferred the supernatant to another fresh tube.
9. Added 900µl of 100% alcohol (Merck), mixed it by vortexing.
10. Kept the tube at room temperature for 5 min.

11. Centrifuged the tubes at 14,000 rpm for 10 min.
12. Supernatant is discarded and precipitate is washed with 70% alcohol.
13. Precipitate is dried and resuspended in TE (10mM Tris.Cl and 1mM EDTA).
14. RNase (Sigma, working concentration 10mg/ml) is added and incubated at 37c for 30 min.
15. After RNase treatment purification of DNA is done by phenol-chloroform method. After that DNA is precipitated by adding 1/10 of 1M NaCl (sigma) and 2.5 volume of absolute alcohol.

2.4 Medi-Prep Method of Plasmid Isolation

Medi preparation of plasmid DNA from 500 ml LB culture was done by alkaline lysis method described by Sambrook et al., (1989).

1. A single colony containing the plasmid (pBluescript) from LB-agar(BDH) plate supplemented with ampicillin (50-100μg/ml) was inoculated in 500 ml of LB medium supplemented with the same antibiotic. The cell were grown overnight (16-18hr) at 300 rpm in 37°c shaker.
2. Cells were harvested by centrifugation at 4000 rpm at 4°c for 10 min.
3. The cells were resuspended in 10 ml of ice-cold **Solution 1** (50mM glucose (Qualigen), 25mM Tris,Cl (Sigma pH 8.0), 10mM EDTA (Sigma)) by vortexing.
4. 20 ml of freshly prepared **Solution 2** (0.2 N NaOH, 1% SDS) is added to the suspension and was mixed gently by inversion and incubated at room temperature for 10 min.
5. 15 ml of 3M sodium acetate pH 5.2 (Qualigen) was added and mixed thoroughly by inverting the tube gently. The tube was incubated in ice for 10 min.
6. The lysate was centrifuged at 12,000 rpm for 15 min, at 4°c and the clear supernatant was carefully transferred to a fresh tube avoiding any precipitate.

7. Nucleic acid was precipitated by the addition of 0.6 volume of isopropanol (Qualigen) at room temperature for 30 min.
8. The nucleic acid was recovered by centrifugation at room temperature at 12,000 rpm. The pellet was washed with 70% ethanol and dried.
9. The pellet was dissolved in TE.
10. After RNase treatment it was extracted once with phenol-chloroform and then with chloroform-isoamyl alcohol. The aqueous phase was precipitated with 2.5 vol of ethanol and 1/10 vol of IM NaCl.
11. The dried pellet was dissolved in TE.
12. Plasmid DNA was purified by PEG precipitation (PEG 8000, Sigma), PEG is added and incubated in ice for 30 min.
13. Plasmid DNA was pelleted at 14,000 rpm at room temperature for 10 min. Washed twice with 70% alcohol. After drying plasmid DNA was dissolved in TE.
14. Estimation of DNA is done spectrophotometrically. The routine yield was 400ug from 500 ml LB grown culture.

2.5 Cloning Of DNA Inserts:-

2.5a Preparation Of Vector And Insert DNA For Cloning

Both fragment and vector digested with appropriate restriction enzymes (HincII [genetix] and SpeI [genetix]), to generate compatible ends for cloning. 10 X Y+/ tango buffer (33mM Tris acetate Ph7.9, 10mM magnesium acetate, 60mM potassium acetate, 0.1mg/ml BSA) is used for maximum enzyme activity. After digestion the DNA of both insert and vector is purified by phenol-chloroform method and then precipitated by 1/10 vol of 1M NaCl and 2.5 vol of ethanol, then it is washed with 70% alcohol .After drying the pellet, it is resuspended in either MQ water or in TE. Quantitation of DNA is done either by spectroscopy or spotting it on agarose (SRL) plate. After that ligation of insert and vector DNA (1:3) is done by T4 DNA ligase (genetix Or Banglore genei) and 5x ligation (400mM Tris.Cl, 100mM $MgCl_2$, 100mM DTT, 5mM ATP (pH 7.8 at 25°c) is used. The reaction is incubated overnight

at 16c. After overnight incubation the ligase is heat inactivated by heating the reaction mix at 65°c for 10 min. After this transformation is done.

2.5b Preparation Of Competent Cells :-

2.5b.I Trituration Method of Competent Cell Preparation:-

DH5αbacterial strain is used for competent cell preparation. DH5α competent cells are made according to the method described by Hanahan (1995) with some modification.

1. DH5α cells were grown on fresh LB-Agar having 20ug/ml nalidixic acid (sigma, stock solution concentration 20 mg/ml) plate from glycerol stock keeping the plate at 37°c overnight.
2. Single colony was inoculated in a preculture of 25ml medium supplemented with same antibiotic.
3. Preculture was grown at 32°c at 300 rpm for 16 hours (primary culture).
4. Inoclated the 100 ml LB media with 1 ml primary culture.
5. Culture is kept for shaking at 300 rpm at 37°c until the OD reaches to 0.45-0.55.
6. Chilled the cells in ice for 2 hours and collected the cell pellet by centrifugation at 4000 rpm for 15 min.
7. Cell pellet is then resuspended in freshly prepared trituration buffer (100 mM $CaCl_2$ (sigma), 70mM $MgCl_2$ (sigma), 40mM sodium acetate (Qualigen), pH 5.5) at equal volumes of the starting culture.
8. Incubated the cells on ice for 45 min.
9. Cells were pelleted by centrifugation at 2500 rpm for 5 min at 4°c.
10. Competent cells were then resuspended in triruration buffer a $1/10^{th}$ volume of the starting culture volume.
11. The competent cells were mixed with sterile glycerol to a final concentration of 15% (v/v).

12. Cells were aliquoted on ice and immediately frozen at -80°c in microcentrifuge tubes and stored at the same temperature until further use.

2.5b.II Na-MOPS Method Ofcompetent Cell Preperation

This method essentially combines from Douglas Hanahan's high efficiency procedure with the traditional calcium-shock methods. This methods provides very high efficiency.

After inoculating the secondary media as I described above the cells were kept for growing.

- Cells were grown until OD reaches to 0.5-0.6.
- Cell pellet was collected after centrifugating the culture at low speed (2500 rpm) for 5 min.
- Drained the cells and resuspended the cells in 40 ml pre 100 ml of starting culture of cold Tfb1 (30mM pot-acetate (Qualigen), 50mM $MnCl_2$ (Qualigen), 100mM KCl (Qualigen), 10mM $CaCl_2$ (sigma), also containing 150 g/l glycerol (Qualigen) final pH 5.0) at 4°c.
- Incubated in ice for 5-10 min.
- Cells were recollected by centrifugation at 2000 rpm for 5 min.
- Resuspended the cell pellet in 4 ml per 100ml starting culture of ice-cold Tfb2 (10mM Na-MOPs (sigma) pH7.0, 75mM $CaCl_2$, 10mM KCl, also containing 150 g/l glycerol, final pH 7.0
- Aliquoted in to tubes and stored at -80°c.

2.5 c. Transformation Of Competent Cells

After competent cell preparation transformation of DH5α is done

1. Competent cells were thawed in ice.
2. Added 1-4 µl (20ng) of ligation mix (ligated vector and insert) to the 100µl cells and mixed it by tapering.

3. Kept in ice for 30 min.
4. The cells were subjected to heat shock at 42°c in a water bath for 5 min. And then chilled on ice for 5min.
5. 900µl of LB is added to the cell after it attained at room temperature.
6. The cells were allowed to outgrow for 1 hour at 37°c to allow the expression of the antibiotic resistance gene.

A positive control of transformation was carried out with 2.5 ng of pBluescript vector and a negative control was kept with out addition of DNA. Transformation efficiency was determined by counting the number of colonies/µg of DNA on selection plate.

2.5d. Plating Of Transformation Mix

20 µl of 40 mg/ml X-gal (genetix) and 100µl of 100mM IPTG (genetix) was spread on 25 ml LB-agar plate containing 100 µg/ml ampicillin. An aliquot of 200 µl out of 1ml transformation mix was spread on each plate. The plates were dried on laminar flow air and then incubated at 37°c incubator, agar side up until the colonies appeared (~ 16 hours). The plates then stored at 4°c.

2.6. Screening Of Recombinants

All the recombinant colony will be of white colour (because recombinants lack β-galactosidase activity) while non recombinants will be of blue colour. Alternatively individual colonies were picked up by a tooth prick, resuspended in 500ul H_2O, centrifuged for 5min at 14000 RPM and one microliter supernatant was used for PCR using T3-T7 or P1-P2 primers.

2.7. Agarose Gel Eelectrophoresis :

Agarose (SRL) gel electrophoresis was done as described in Sambrook et al., 1989.

Routinely 0.5 X TBE buffer (0.9 M Tris (SIGMA), 0.9 M Boric acid (SIGMA), 0.025 M $Na_2EDTA,2H_2O$(SIGMA), autoclaved, pH 8.0) was used Ethidium bromide (BDH) was added to 0.5 μg/ml concentration in gel. Generally, 0.8 % gels and 2% gels were used.

2.8. Acrylamide Gel Electrophoresis

30 % acrylamide solution (29 gm acrylamide (sigma) and 1 gm N-N' Bis acrylamide (sigma) made it 100 ml) was prepared. Generally 8% gel was used.

For 50 ml gel-

10 X TBE	2.5 ml
30% Acryl amide	13.3 ml
10% APS (Amresco)	200μl
TEMED (Amresco or SRL)	40 μl.
H_2O	To 50 ml

2.9. Elution of DNA

The fragment of was cut from the gel and eluted. For elution the gel fragment / piece containing DNA or DNA- protein complexes was chopped, put in a tube and elution buffer (500mM, Ammonium acetate, 10mM $MgCl_2$, 1mM EDTA, and 0.1% SDS) was added in the tube. The tube was kept overnight on shaker. After 20 hours the supernatant was transferred in to another tube. In the tube 1/10 NaCl and 2.5 vol alcohol was added for DNA precipitation, and kept for 2 hr in -20°c. After that it was centrifuged at 12,000 rpm for 15 min, washed twice with alcohol, dried, and resuspended in TE.

2.10 Electrophoresis

The gel was run at 200 V for 3 hours. Before loading the samples the lanes of the gel were flushed with the needle of a syringe to prevent the uneven

loading. For gel shift assays the gel was run at 250V for 3 hour in 0.5X TBE buffer. After complete run the gel was stained with ethidium bromide solution.

2.11 Probe Preparation

Probe for gelmobility shift assay was done by PCR using recombinant plasmids as template. The parameters were same as the routine PCR except the nucletide (dNTP minus dCTP) concentration was 10mM and 20uCi of αP^{32} dCTP was used in 20 ul reaction. The reaction product was purified by filtration through Sephadex G 25 column.

2.12. Preparation of Nuclear Extracts

13 days old chick embryo is taken and nuclear extract was prepared according to Dignam et al. All steps were conducted at 0-4°c unless stated otherwise.

1. Brain, heart, liver and skeletal tissue were removed and added to the prechilled hand homogenizer.
2. One ml of **buffer A** (20mM Hepes (sigma), 20% glycerol (Qualigen), 10mM NaCl (sigma), 1.5mM MgCl2 (sigma), 0.2mM EDTA (sigma) 0.1% triton-X100 (sigma), 100mM PMSF (sigma), 10mg/ml Leupeptin (sigma), 10 mg/ml Aprotenin (sigma)), is added in each homogenizer having individual tissue.
3. The tissues were homogenized.
4. The homogenate is centrifuged at 2500 rpm at 4°c for 15 min.
5. Supernatant is discarded, pellet is then resuspended in 200-300 ul of **buffer B** (20mM Hepes, 20% glycerol, 500mM NaCl, 1.5mM $MgCl_2$, 0.2mM EDTA, 0.1% TRITON-X, 100mM PMSF, 10mg/ml Leupeptin, 10mg/ml Aprotenin).
6. Incubated the extract on ice for 1 hour with tapping at every 10 minutes interval.
7. The extract is then centrifuged at 15,000 rpm at 4°c for 15 min.

8. The supernatant is aliquoted in to small tubes, and stored the extract rapidly at -70°c.

2.13 Protein-DNA Binding Assay

Binding reaction with tissue extract were performed in a volume of 40 µl. Binding reaction contain 5X GSA buffer (100mM Hepes, 7.5mM $MgCl_2$, 5mM DTT, 20% glycerol). In a typical binding reaction 20-04 ug of nuclear protein is incubated with 1ug of polydIdC, 5X GSA buffer (8ul/40ul reaction), 1mM EDTA and radiolabeled probe (20,000-40,000 CPM). All the compounds of the binding reaction except probe were mixed in a tube on ice. Probe DNA is then added. The reaction is then kept for incubation in ice for 1 hour. DNA-protein complexes were separated from the free probe by electrophoresis in 6% non-denaturing acrylamide gel. Gel electrophoresis continued until the bromophenol dye in the gel lane migrate through 80% of the gel length. The gel was then transferred to a whattmann 3mm paper, dried and exposed to a film. For competition unlabeled DNA was added to the binding reaction prior to the addition of the labeled probe. Competition were done with a 50 fold molar excess over the probe amount.

2.14. Fixation and Autoradiography

After the gel run, plates were separated with the help of spatula. The gel is transferred to the Whatmann 3mm paper and dried in gel dryer at 75°c for 45 min. The dried gel was exposed to the XAR (KODAK) X-ray film and kept inside the cassette. The cassette is kept over night at -70°c. after that the film was developed by developer (KODAK) and fixed in fixer (KODAK).

2.15. DNA Sequencing

Took the DNA which had been purified by PEG precipitation.

Dissolved it in 10ul MQ and quantitated, (for 3kb template size 200ng DNA was used).

Cyclic sequencing is done for that reagents were thawed in ice.

Combined the following components (total volume of 17μl) in a 0.5-ml microcentrifuge tube labeled "premix".

15pmole of unlabelled T3 primer

10μCi of [α-^{35}S]-dATP at 1,000Ci/mmole

7.2 μl of Sequi Therm EXCEL II Sequencing buffer (Epicenter Biotechnologies, USA)

200 ng DNA template (3 Kb)

deionized water to 16ul

1μl of Sequi therm EXCEL II DNA Polymerase (5U/μl, Epicenter Biotechnologies, USA)

For each template, four 0.5-ml microfuge tubes were labeled as G, A,T, C and placed on ice. Then added-

2 μl of SequiTherm EXCEL II Termination mix to the G tube

2 μl of Sequi Therm EXCEL II Termination mix to the A tube

2 μl of Sequi Therm EXCEL II Termination mix to the T tube

2 μl of Sequi Therm EXCEL II Termination mix to the C tube

On ice, 4μl of the premix was added to each of the four tubes and mixed thoroughly.

Overlaid each reaction with mineral oil and the tubes were centrifuged briefly in a microcentrifuge to separate the mineral oil layer from the reaction component.

Preheated the thermocycler to 95°c.

Heated the reactions for 5min at 95°c.

Cycled the reactions 30X for

30sec at 95°c

30sec at 55°c

1min at 70°c

Upon completion 3 µl of stop/loading buffer was added to each reaction, and proceeded with electrophoresis.

2.16. Gel Electrophoresis for Sequencing

Tubes were centrifuged to separate mineral oil from the reaction.

6% acrylamide-8M urea gel was prepared as follows

For 70ml solution

Urea	35g (8M)
40% acrylamide	10.5ml(6%)
10X TBE	7ml
10% APS	0.7ml
TEMED	13.3 µl

Before loading the tubes were heated at 95°c.

After the complete gel run, it was treated with the fixer

20 % fixer (1000 ml)

100 ml methanol

100 ml acetic acid

800 ml double distilled water

Gel is then dried and exposed to phosphoimager.

2.17. Prime sequences

T7	5' TAATACGACTCACTATA 3'
T3	5' AATTAACCCTCACTAAA 3'
P1	5' CGAGGTCGACGGTATCG 3'
P2	5' GCTCTAGAACTAGTGGATC 3'
M1	5' CGAGGTCGACGGTATCG 3'
M2	5' GCTCTAGAACTAGTGGATC 3'

Aims and Objectives

AIMS AND OBJECTIVES:

The present study is aimed at identifying novel transcription factor activities from developing embryo and to evaluate their role in regulating tissue specific gene expression. Conventional approach to such problem is (i) to identify a gene(s) that is either tissue restricted or developmentally regulated, (ii) isolate its promoter and identify the cis-regulatory elements (iii) identify, characterize and clone the cognate transcription factors. Such approach is labor intensive and is inherently biased on the gene being taken as the subject of study. To circumvent these limitations, researchers have attempted gel mobility shift assay in conjunction with randomized target site binding selection to find novel transcription factor activities (120). However, such methods are yet to be perfected for general application. Our laboratory has significantly modified this technique and is currently using it to identify novel transcription factor activities from developing chick heart. In the present study I have adopted this strategy to identify novel transcription factor activities from various organs in the developing chick embryo. To achieve this goal, the following work has been done:

1) Construction of a library of randomized, ten nucleotide long DNA sequences in the plasmid vector pBluescript.
2) Screening of the library with nuclear extracts from the liver, brain and heart from twelve-day chick embryo to select a pool of target sites.
3) Identification of several of those binding activities which are differentially expressed in various tissues.
4) Cloning the selected sites into pBluescript vector.
5) Identification and primary characterization the individual binding sequences.

Results

2.a Construction of a Library of Random Target Sites

The primary objective of this study is to isolate novel tissue restricted transcription factor activities from various organs of developing chick embryo. To that purpose, a synthetic 47 nucleotide long oligonucleotide harboring a randomized ten nucleotide long core sequence flanked by a number of restriction sites (Fig.1A) was commercially synthesized (Genosys, USA). The single stranded synthetic DNA was then converted into the double stranded form by Polymerase Chain Reaction using a primer pair corresponding to the terminal sequences as shown in Fig 1A. The amplified product was then digested by restriction enzymes HincII and SpeI (Fig 1B) and ligated to the Hinc II-SpeI linearized plasmid vector pBluescript (Fig 2A & B). The ligation mixture was transformed into E.coli strain DH5α competent cells. Normally 10ul of ligation mixture was transformed into 200ul competent cells and the entire mixture was plated on a 150 mm petridish containing LB agar medium with 100ug/ml ampicillin. From one such transformation, about one thousand penicillin resistant colonies were obtained. In order to determine the recombinant frequency, twenty colonies were randomly picked up, resuspended in water, heat denatured, and the multiple cloning sites of the harboring plasmids were amplified by the T3-T7 primer pair. For the non-recombinants (vector) the amplification product were 149bp while that for the recombinants, it was 131bp. Data from a representative experiment for determining the recombination frequency is shown in Fig 3. This particular batch had 100% recombinants. Nonetheless, the recombination frequency was generally 80% or above and only those which contained at least 80% recombinants were taken for subsequent study. Following such initial characterization of each batch of the library, the colonies from each plate were pooled and collected as cell pellets. About 10,000 such recombinants (11,000-12,000 total colonies) were finally pooled together and saved as one set of library. Since the target size for transcription factors vary from four-nucleotide sequence (GATA) to as large as nine nucleotide (MEF-2, TTATTTTTA) or more, the chance of finding a novel four nucleotide sequence is one in 256 (1×4^4) recombinants while that of a seven nucleotide

sequence is one in 16384 (1 X 4^9). We thus rationalized that, if the number of target sites for such selection is very large (50,000 or more), the chance that one single tetra nucleotide binding site will occur 200 times and will thus be over represented in the selected pool as compared to those which consists of larger sequences. Furthermore, taking a large number of recombinants in a single pool may also diminish the abundance of individual sites and as a result, may compromise the chance of binding relatively less abundant factors. With that rationale, we opted for selecting the target sites from a pool of twenty thousand recombinants in each batch. To that purpose, two such pools of recombinants were combined and the plasmid DNA isolated from the mixture was used for the selection study as described below.

2.b. Selection of Transcription Factor Target Sites by Different Tissue Extracts:

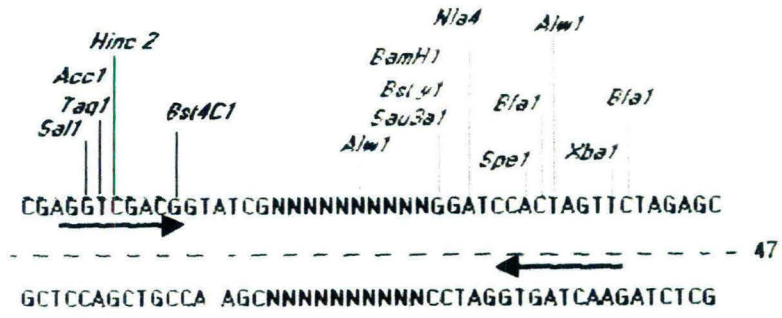
Following construction of the library of putative cis elements as described above, a pool of about 20,000 recombinants were taken and the harboring plasmids were isolated following standard protocols. The T3-T7 primer pair was used for amplifying about 100 picogram of the recombinant plasmid pool. The quality and the identity of the amplified product were confirmed by running on 8% acrylamide gel (Fig 4). The amplified product was predominantly the 131 bp recombinant and a small proportion was that derived from the accompanying vector (149 bp, not visible in Fig 4). Ten nanogram of the purified, amplified product was then used for gel mobility shift assay using nuclear extract from liver, brain, heart and cardiac tissue isolated from 12 day-old chick embryo. The strategy for this selection process is described Fig 5. It essentially consists of the following steps. The amplified pool of sequences were allowed to bind to the nuclear proteins under optimum condition and the bound sequences were separated from the unbound one on 8% native acrylamide gel. The entire length of gel containing the protein bound sequences were eluted, purified, amplified and reselected. The quality of the nuclear extracts were first confirmed by running a control gel shift assay with radio-labeled sp-1 oligonucleotide (Fig 6). For the selection reaction, a parallel run

was done with 100ng target sequence as a marker and with that as reference, the entire length of the containing the protein bound sequences were excised (see Fig 7 as a representative experiment). The retarded DNA sequences were eluted and reamplified by a set of internal primer pairs and the selection process was repeated three times. We observed that following elution, the subsequent amplification is prone to generate a strong smeary background. To circumvent this problem, in each amplification step, following elution a new set of internal primer pairs were used (Fig 8). Following three consecutive cycles of selection-amplification, the eluted material was finally amplified by the innermost primer pair, digested by Spe-1-Hinc II and cloned again into the pBluescript vector. Among the candidate clones, those harboring the recombinant were identified by PCR as described before (Fig 9), and twenty of such recombinant pool was taken for the subsequent study.

2.c. Characterization of the Selected Target Sites:

Following three independent selection of the target sites with brain liver and heart nuclear extracts, the binding profile of those selected sequences were compared. Twenty recombinants were randomly picked up from each batch and were processed for plasmid isolation. The purified plasmid was used as a template for the preparation of radio-labeled probe by PCR using the innermost primer pair and α-^{32}P-dCTP. Three such probes (brain, liver and heart select) were then independently tested by gel mobility shift assay using extracts from 12 day old brain, heart, liver and skeletal muscle. As shown in Fig 10, each probe generated a similar profile except a few distinct differences. For example, each probe showed one liver specific/enriched (Fig 10, complex L) complex and one heart specific/enriched complex (Complex H) irrespective of the source of the extract for its selection. The binding profile thus clearly established the efficacy of this method for selecting DNA sequences targeted by sequence specific DNA binding proteins. In order to further establish this protocol, three distinct complexes (C1-3, Fig 10), were identified, excised from the dried gel after aligning with the autoradiogram. The bound DNA was eluted, amplified by the M1/M2 primer pair and cloned into pBluscript vector. For each eluted material, a probe was also generated by PCR

amplification of the template and was used for gel shift assay with respective nuclear extracts. Each such probe tested generated a distinct binding profile that was competed out by the unlabelled competitors produced from individual clones by PCR amplification of the insert using T3/T7 primer pair (Fig 11 A, B, & C). In this competition assay, specific complexes were competed out by majority of the competitors and thus demonstrated the specificity of the binding activity. Based on such competition assays, three independent competitors for the complex 3 were selected and sequenced. The sequence of the respective insert is shown in Fig 12.



`**Fig1A**. Restriction map of the forty-seven nucleotide long synthetic oligonucleotide used for target site selection. The ten-nucleotide long randomized core sequence is identified by N. The primar pairs used for its amplification is also identified.

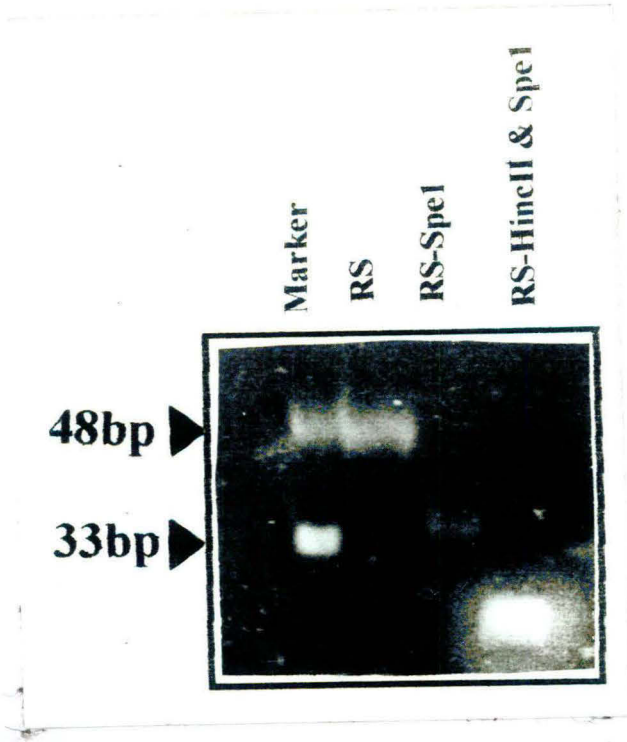


Fig1B. Restriction digestion of 47 bp oligonuleotide with Spe I and HincII. Lane 1: Marker (multiple cloning site of pBluescript amplified with M1/M2 primer pair and digested with EcoR1). Lanes 2: undigested DNA, lane 3: Spe1 digest, Lane 4: HincII & Spe1 digest.

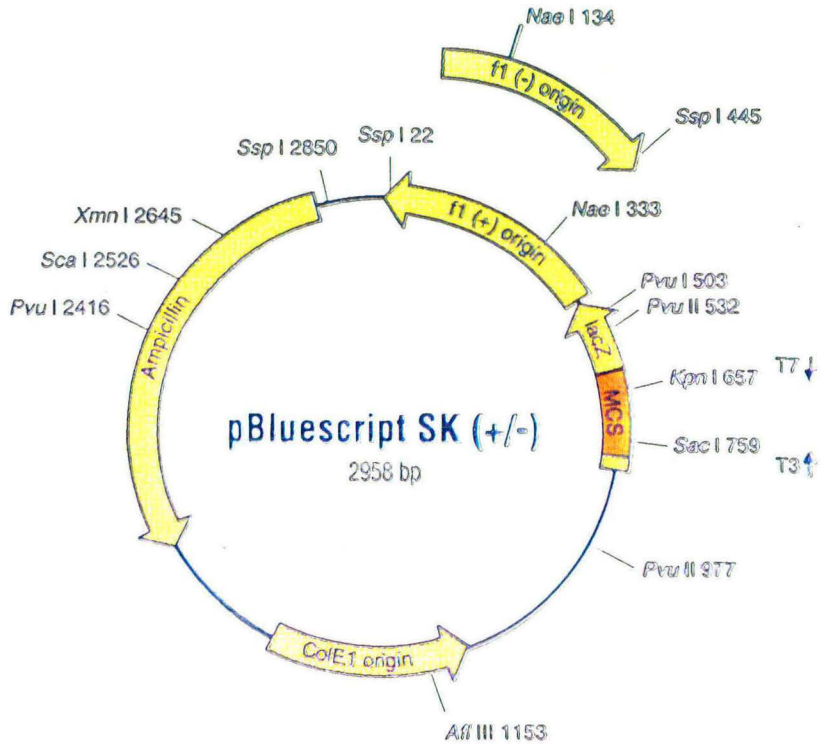
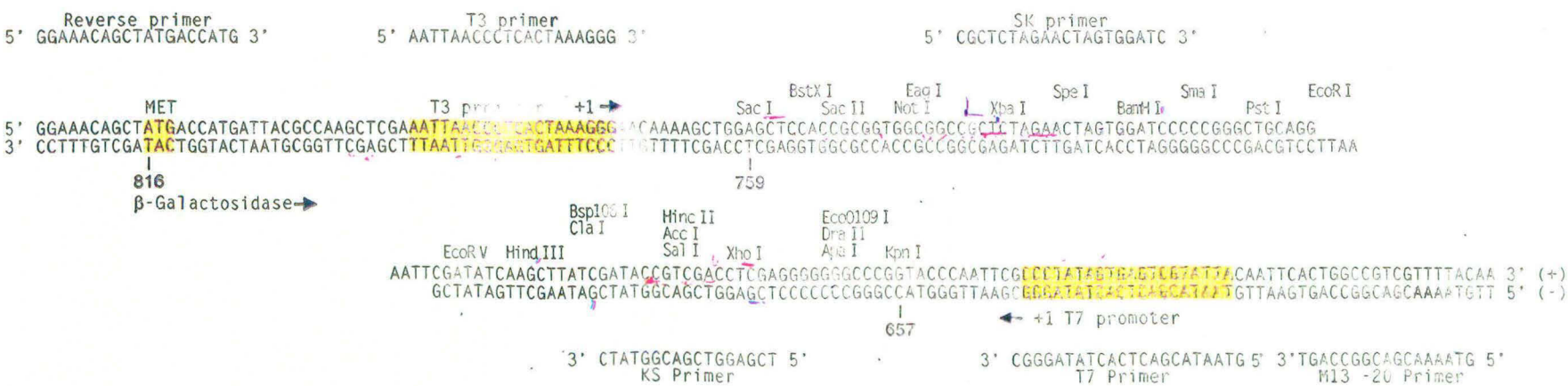


Fig2A Restriction map of the multiple cloning site of the pBluescript vector flanked by the T3/T7 primer pair.



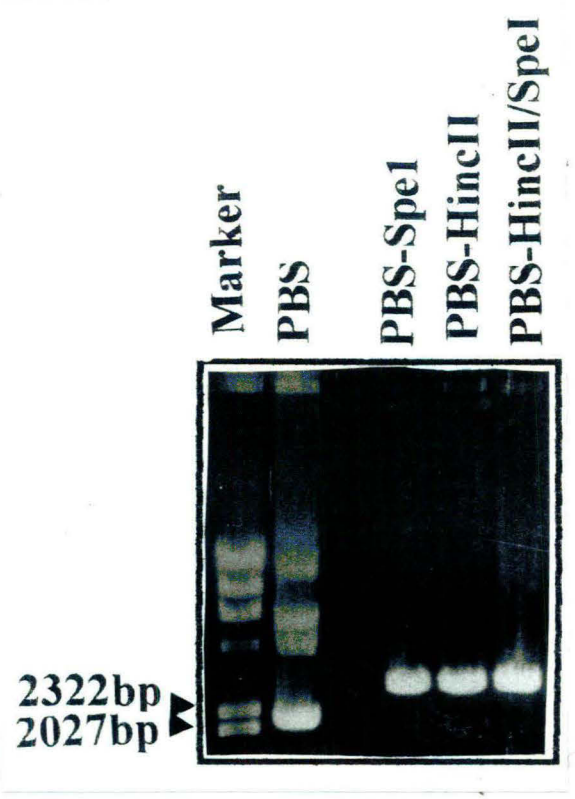


Fig2B. Restiction digestion of pBluescript with SpeI and HincII. Lane 1: size marker (ϕx174 DNA digested with Hae III), lane 2: undigested pBluescript plasmid. lanes 3-5: pBluescript digested with SpeI, HincII and **both** SpeI & HincII respectively.

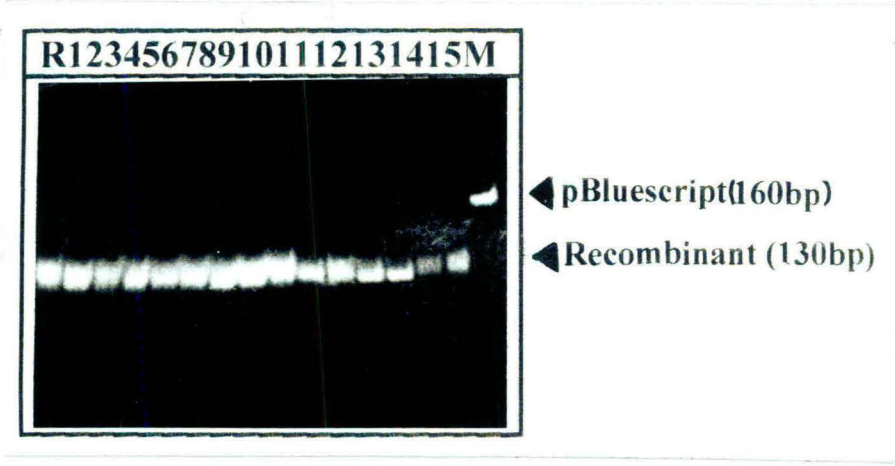


Fig3. Determination of the recombinant frequency of a representative a library. Fifteen colonies were randomly picked up, heat denatured, amplified with T3/T7 primer pair and checked on 8/% acrylamide gel. Lanes 1-15 represent recombinant bands and lane M indicates pBluescript multiple cloning site amplified with T3/T7 primer pair. In this particular batch all the cloned tested were recombinants.

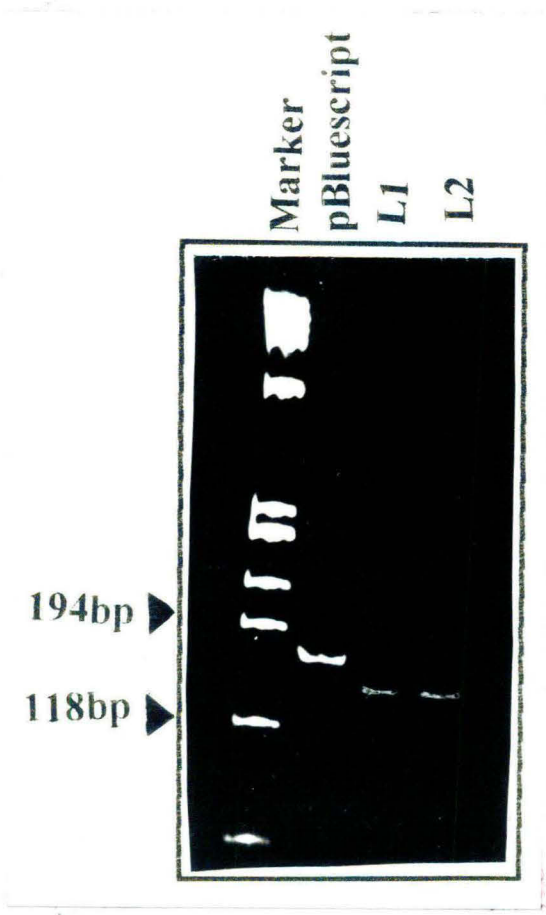


Fig4. Pool of pBluescript DNA harbouring 10 nucleotide random sequence were amplified by T3/T7 primer pair. The integrity of the product was checked on 8% acrylamide gel. Lane 1(HaeIII digested ϕx 174 DNA), lane 2: pBluescript multiple cloning site amplified with T3/T7 primer pair. Lanes 3& 4: two independent libraries L1 and L2, each consisting of 10,000-12,000 recombinants, were amplified by T3/T7 primer pair. The 130bp long recombinant product is shown. The 149 bp long product from contaminating vector is not visible.

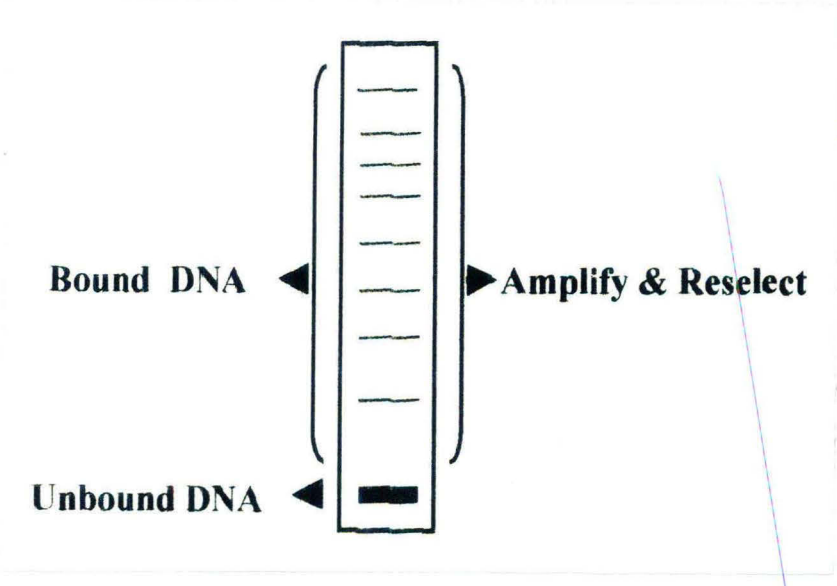


Fig. 5 The strategy for target site selection is shown. The DNA fragments bound to their cognate proteins are separated on acrylamide gel and eluted for further amplification. Repetition of this process leads to a pool of selected sequences, which are subsequently cloned in to the vector for further characterization.

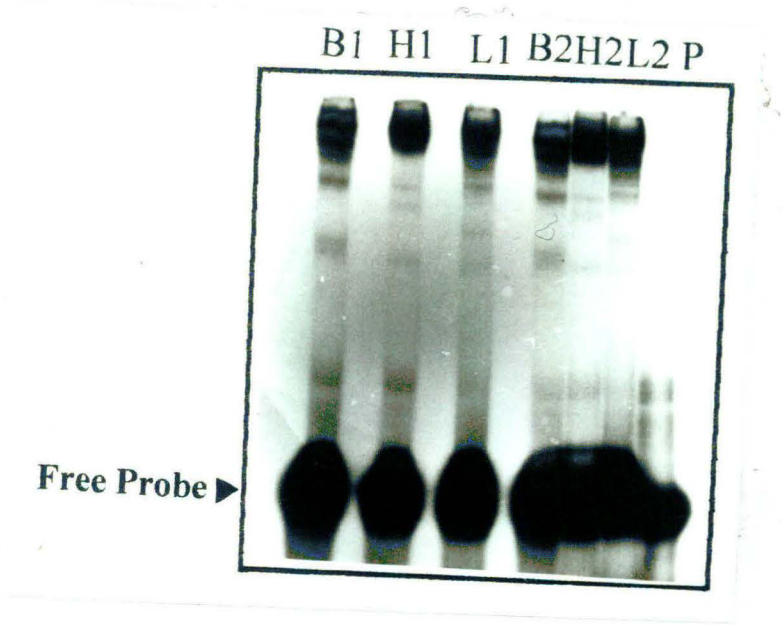


Fig6. Integrity of nuclear extracts prepared from various tissues were checked by gel mobility shift assay with radiolabeled sp-1. Lane P represents the free probe. B: brain extract, H: Heart extract, L: Liver extract. Two independent batches were tested for subsequent use.

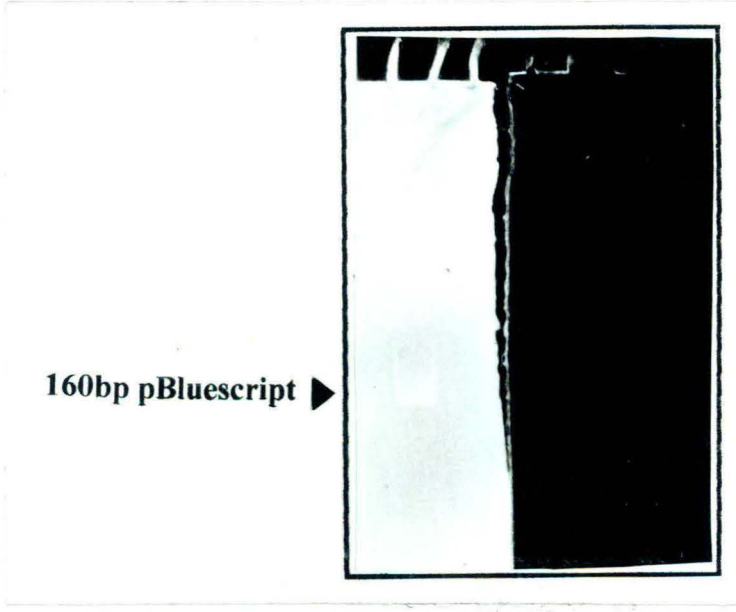


Fig7. Initial step for isolating probable binding sequences by electrophoretic mobility shift assay. Ethidium bromide stained portion of the gel containing the marker (pBluescript multiple cloning site amplified by T3-T7 primer pair was aligned with the unstained portion of the gel containing the DNA-protein complex. The entire lane starting from the well till 1cm above the marker was excised and eluted.

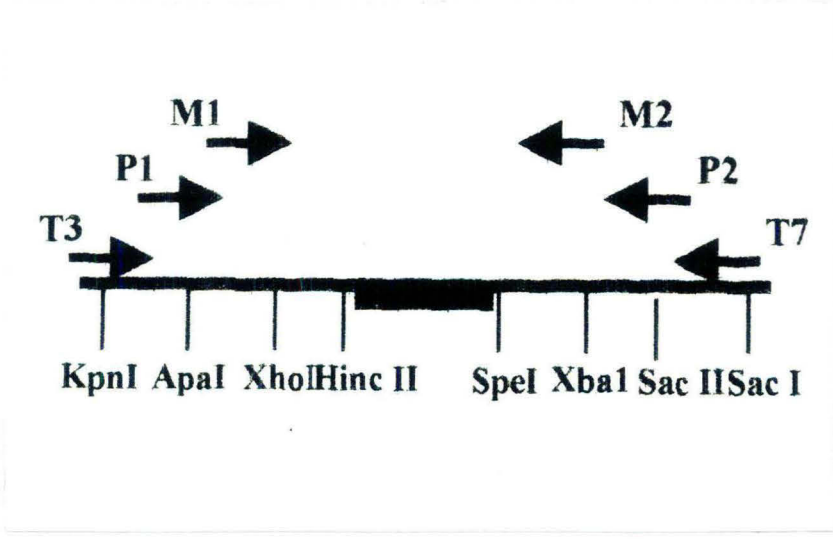


Fig8. Three different primer pairs used for the selection-amplification procedure. Each batch of eluted material was amplified by pair of internal primers to eliminate the spurious amplification products.

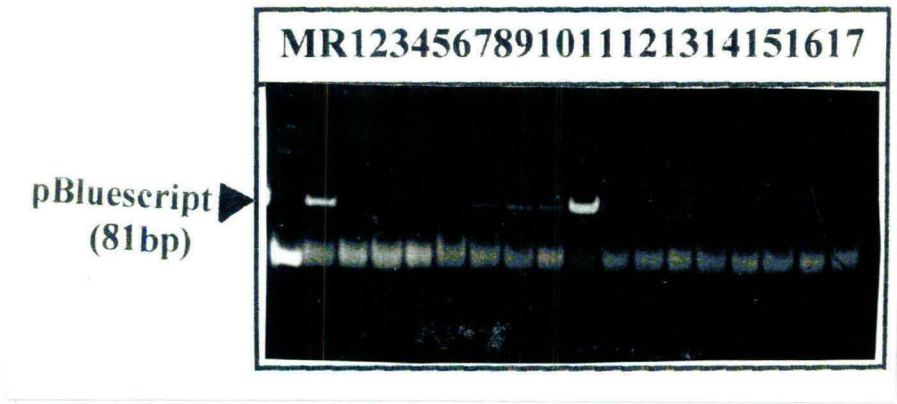


Fig9. Determination of recombination frequency after selection.After four rounds of selection, the selected DNA was cloned in pBluescript and tested for recombination frequency. Lane1: the marker, Lane 1-17: different clones. Presumably clone 1,6,7,8,& 9 are nonrecombinants, possibly conteminated by recombinants.

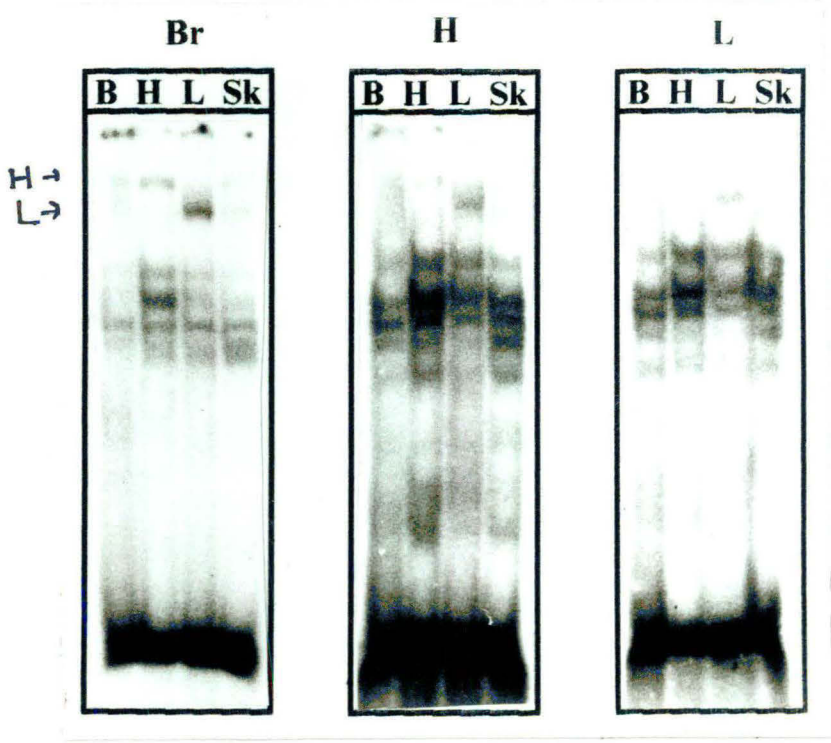


Fig10. Tissue distribution profile of the binding activities of the selected target sites. Pool of plasmid DNA (20,000) harbouring putative target sequences were independently selected by brain, heart and liver extracts from twelve day old chick embryo. Following selection, the sequences were cloned and twenty recombinants from each pool was tested for binding activity. Probes were generated from each selected pool . Profile of binding activities by those selected sequences were tested using extracts from twelve day old brain, heart, liver and skeletal muscle tissues.

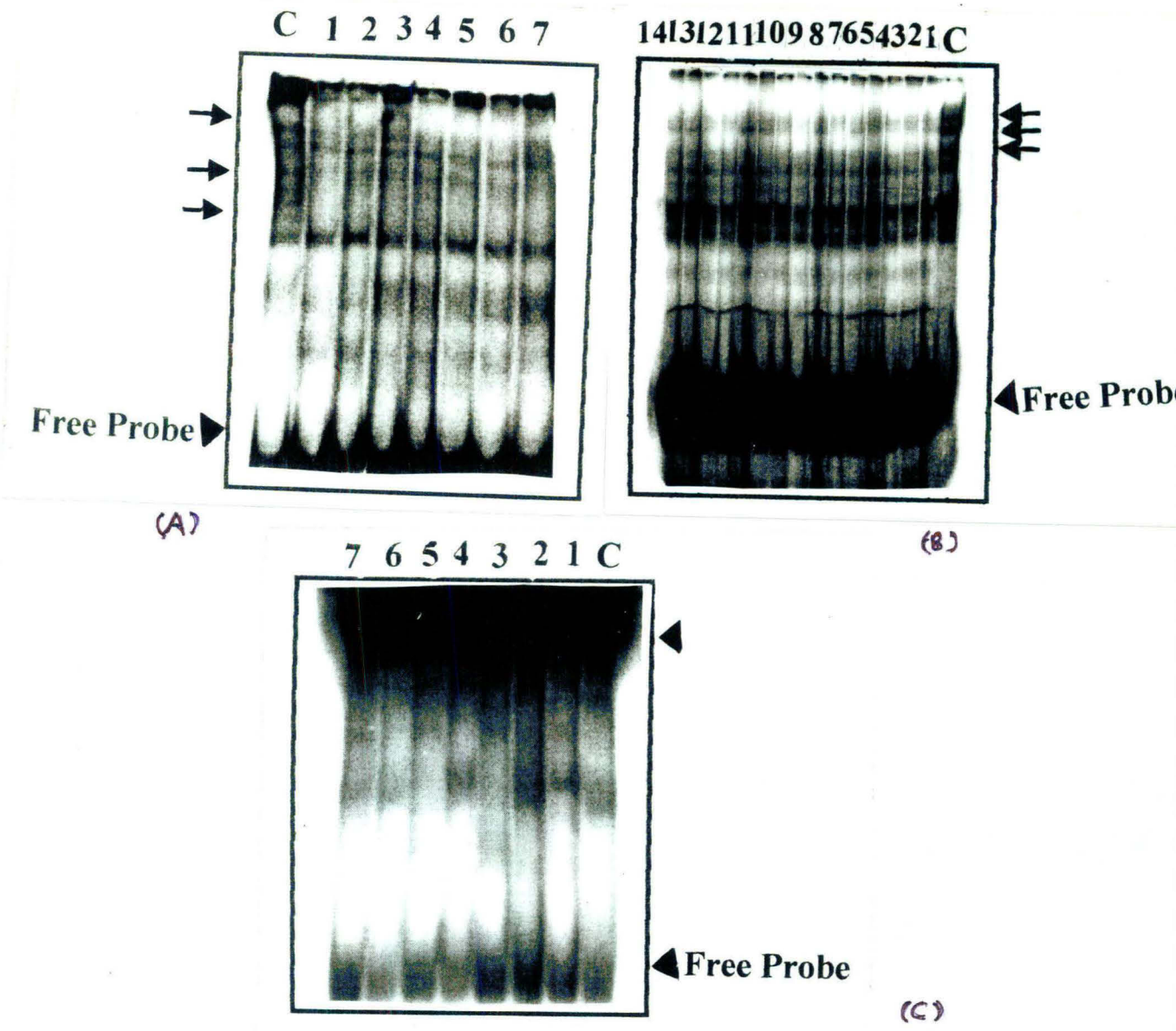


Fig11. Differential binding profile of candidate target sites. One candidate clone from each tissue (brain, heart and liver) select was excised from the dried gel, eluted out and amplified. The amplified products were used as templates to generate radiolabeled probe. Also the amplified product of each was cloned into vector and individual clones were used for competition assay. A. Brain select, B:Liver select, C: Heart select. In each Fig, lane 'C' represents binding reaction of the probe without competitor.

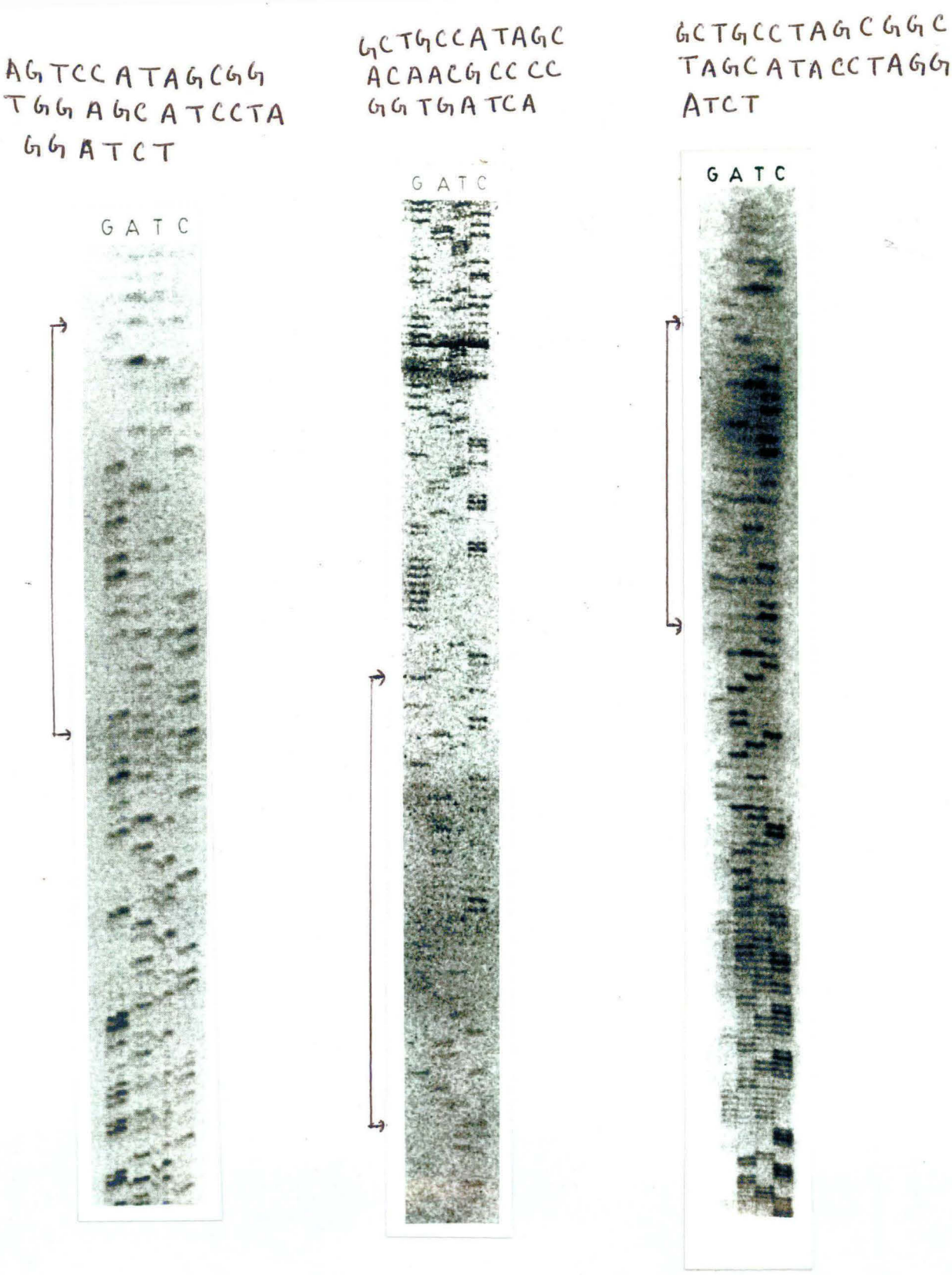


Fig12. Cyclic sequencing of pBluescript multiple cloning site harbouring 10nucleotide long randomized sequence.

Discussion

DISCUSSION

Differential gene activities underlie the basic mechanism of disparate biological processes such as cell differentiation, organ development, tumor growth etc. Among various steps in the regulation of gene expression, control of transcription plays a major role in tissue specific gene expression. During embryogenesis, following receiving stimulatory cues from the surrounding population of cells, a distinct group of embryonic cells initiates a distinct transcriptional program that leads to their differentiation. Discovery of MyoD family of transcriptional regulators has significantly enhanced our present understanding of molecular basis of cell differentiation (). However, studies done during the past several years, on the differentiation of other cell types, have revealed that the initiation of skeletal muscle cell differentiation by a small group of regulators is perhaps the exception not the rule. It is believed that regulation of differentiation of other cell types is controlled by a complex combinatorial effect of a number of regulators rather than a single master switch (see introduction for details). It is therefore imperative that identification of the desperate family of transcriptional regulators from various organs is a prerequisite for understanding the molecular basis of differentiation of various cell types. The conventional way of identifying a novel transcriptional regulator is to characterize the promoter region of a representative gene, identify various cis regulatory elements that control it expression and characterize the cognate factors. However this process is labor intensive and is inherently biased for the selection of the representative gene. To circumvent this problem, molecular biologists over the years have attempted to develop methods for the identification of novel genes from differentiating cells by methods such as differential display, subtractive hybridization. However these methods may lead to the identification of any gene rather than those specifically involved in regulating the differentiation program. On the contrary, PCR based methodologies for the selection of specific DNA sequences targeted by transcription factors have been attempted by researchers for selectively identifying the transcription regulators involved in this process (). Our laboratory has significantly improved this methodology and have established it as an effective method for searching novel transcription factors from

any cell type. One major feature of this modification is, instead of using the entire population of the target sequences, a small fraction are cloned into a plasmid vector which enhances the kinetic of interaction between the target site and the binding protein. In the present study one such small pool (20,000) of target sequences were first cloned into plasmid vector and then used for the selection using nuclear extracts from brain, liver and heart. Following three rounds of such selection, the selected sequences wee again cloned into the vector and twenty independent clones were tested for their binding activity. The binding profile thus generated was very similar among each of those selected pools as expected. In eukaryotic cell majority of the transcription factors are either ubiquitous or at least different members of the same family bind to the same target sequence. However, I observed that one specific complex (L) in that profile was liver enriched, irrespective of whether it was selected by liver extract or not. It means that the selection process used in the study is sensitive enough to pick up very low abundant activities that is not normally visible in the gel shift assay. Selection of another such sequence (H) that is enriched in heart but also could be selected by brain extract also supports this conclusion. Following successful isolation of those binding sites, three representative sequences, one each from the brain, liver and heart select were partially characterized. Gel mobility profile with those candidate sequences, in conjunction with the competition assay; clearly demonstrate their binding specificity. Sequence analysis of three candidate clones indicate that perhaps the targeted sequence is degenerate in nature and thus requires further characterization by other assays such as DNAse I foot printing. Therefore, it is established that distinct DNA sequences targeted by transciption factors can be isolated from a mixture of random sequences. This study has thus established the efficacy of this method for the profiling transcription factor activities from various tissues and has opened avenues for future isolation of tissue restricted and developmental stage restricted transcription factor activities.

Conclusion

CONCLUSIONS

A method has been established for the efficient isolation and characterization of DNA sequences targeted by various transcription factors. This strategy is sensitive enough to pick up sequences which are targeted by very low abundant and tissue restricted factors. A number of DNA sequences thus isolated show distinct binding profile in gel mobility shift assay and thus open the avenue for future isolation of tissue/developmental stage restricted transcription factor activities.

References

1. Anthony J. Koleske and Richerd A.Young; The RNA Polymerase II holoenzyme and its implications for gene regulation TIBS 20,113-116, March 1995.

2. George Orphanides, Thierry Lagrange, and Danny Reinberg: The general transcription factors of RNA plymerase II;Genes and Development 10, 2657-2683, 1996.

3. David S. Latchman;Transcription Factors: An Overview ;Int. J Biochem. Cell Biol 29, 1305-1312, 1997.

4. Carl O. Pabo, Robert T. Sauer: Structural Families and Principles of DNA Recognition; Annu. Rev. Biochem. 61, 1053-95, 1992.

5. Robert G. Roeder: The role of general initiation factors in transcription by RNA polymerase II; TBIS 327- 335, 1996.

6. Arnold J Berk: Activation of RNA Polymerase II transcription Current Opinion in Cell Biology; 11, 330-335, 1999.

7. Roger D. Kornberg:Eukaryotic transcription control; TCB 9 (12), M46-M49, 1999.

8. E Peter Geiduschek and George A Kassavetis: Comparing transcriptional initiation by RNA polymerases I and III; Current Opinion in Cell Biology 7, 344-351, 1995.

9. Peter R. Cook: The Organization of Replication and Transcription; Science 284, 1790-1795, 1999.

10. Alan K. Kutach and James T. Kadonaga: The Downstream Promoter Element DPE Appears To Be as Widely Used as the TATA Box in Drosophila Core Promoters; Molecular and Cellular Biology 11, 4754-4764. 2000.

11. Peter Verrijzer and Robert Tjian; TAFs mediate transcriptional activation and promoter selectivity; TIBS 21-September, 338-345, 1996.

12. Frank Sauer and Robert Tjian; Mechanisms of transcriptional activation: differences and similarities between yeast, Drosophila, and man; Curr.Opin. Genet. & Dev.; 7, 176-181, 1997.

13. Yutaka Hirose and James L. Manle: RNA Polymerase II and the integration of nuclear events; Genes & Development 14, 1415-1429, year

14. Stephen Buratowski: The Basics of Basal Transcription by RNA Polymerase II ; Cell 77, 1-3, 1994.

15. Robert Tjian and Tom Maniatis: Transcriptional Activation: A Complex Puzzle with Few Easy Pieces; Cell 77, 5-8, 1994.

16. Peter Seither, Sebastian Iben and Ingrid Grummt.: Mammalian RNA Polymerase I Exists as a Holoenzyme with Associated Basal Transcription Factors; J. Mol. Biol. 275, 43-53, 1998.

17. Mark Ptashne & Alexander Gann: Transcriptional activation by recruitment; Nature 3861, 569-577, 1997.

18. Christopher C. Adams and Jerry L. Workman: Nucleosome Displacement in Transcription; Cell 72, 305-308, 1993.

19. Charlotte A. Spencer & Mark Groudine: Transcription elongation and eukaryotic gene regulation; Received 8 March 1990: Accepted 9 March 1990.

20. Wendy Hanna - Rose and Ulla Hansen: Active repression mechanisms of eukaryotic transcription repressors; TIG 12 (6), 229-234,1996.

21. Stig K. Hansen, Shinko Takada, Raymond H. Jacobson, John T, Lis: Transcription Properties of a Cell Type - Specific TATA - Binding protein, TRF; Cell 91, 71-83, 1997.

22. Peter C .van der Vliet and C. Peter Verrijer: Bending of DNA by transcription factors; Bio Essays 15 (1), 25-31, 1993.

23. Stephen P. Bell, Craig S. Pikaard, Ron H. Reeder, and Robert Tjian: Molecular mechanism governing species-specific transcription of Ribosomal RNA; Cell 59, 489-497, 1994.

24. David C. Ludolph and Stephen F. Konieczny: Transcription factor families : muscling in on the myogenic program; FASEB J. 9, 1595-1604 ,1995

25. Gary E Lyons: Vertebrate heart development; Curr. Opin. Genet. & dev. 6, 454-460, 1996.

26. Akira Ishihama: Protein-Protein Communication within the Transcription apparatus; Journal of bacteriology ,2483- 2489, 1993.

27. C K Glass: Some new twists in the regulation of gene expression of gene expression by thyroid hormone and retinoic acid receptors; Journal of Endocrinology 150, 349-357, 1996.

28. Jean-Yves Nothias,Sadan Majumder,Kotaro J.Kaneko,and Melvin L.DePamphilis: Regulation of gene expression at the beginning of mammalian developmen; The journal of Biological Chemistry 270(38), 22077-22080,1995.

29. Mattias Mannervik,Yutaka Nibu,Hailan Zhang,Michael Levine: Transcriptional coregulators in development; Science 284, 606-609, 1999.

30. Kyuson Yun and Barbara Wold: Skeltal muscle determination and differentiation: story of a core regulatory network and its context; Current Opinion in Cell Biology 8, 877-889, 1996.

31. Eric N.Olson ,Michael Perry,and Robert A.Schulz: Regulation of muscle differenciation by the MEF2 family of MADS box transcription factors; Developmental Biology 172, 2-14, 1995.

32. Hans-Henning Arnold and Barbara Winter: Muscle differentiation:more complexity to the network of myogenic regulators; Curr. Opin. Genet. & Development 8, 539-544, 1998.

33. Brigid L.M.Hogan: Bone morphogenetic proteins: multifunctional regulators of vertebrate development Genes &Development 10, 1580-1594,1996.

34. Harnold Weintraub: The MyoD family and myogenesis: Redundency, networks and thresholds; Cell 75, 14241-1244, 1993.

35. Margaret Buckingham: Molecular biology of muscle development; Cell, 78,15-21,1994.

36. Eric N. Olson and Nadia Rosenthal: Homeobox genes and muscle patterning; cell 79, 9-12,1994.

37. Eric N.Olson and william H.Klein: bHLH factors in muscle development:deadlines and commitments,what to leave in and what to leave out; Genes & Development 8:1-8,1994.

38. Eric N. Olson: Regulation of Muscle transcription by MyoD family; The heart of the matter ; Circulation research 72:1-6,1993.

39. DIanne G.Edmondson and Eric N.Olson: Helix-Loop-helix proteins as regulators of muscle -specific transcription; The Journal of biological Chemistry 268 (2), 755-758,1993.

40. Eric n. Olson: MyoD family: a paradigm for development?; Genes & Development 4,1454-1461,1990.

41. Brian K.Hall and T.Miyake: All for one and one for all: Condensation and the initiation of skeletal development; BioEssays 22, 138-147,2000.

42. Henry F.Epstein and Sanford I.Bernstein: Genetic approaches to understanding muscle development; Developmental Biology154,231-244,1992.

43. N.Rosenthal: Muscle cell differentiation; Current Opinion in Cell Biology 1, 1094-1101, 1989.

44. Randy L.Jhonson and Clifford J.Tabin: Molecular models for vertebrate limb development; Cell 90, 979-990,1997.

45. Giulio Cossu, Shahragim Tajbakhsh and Margaret Buckingham: How is myogenesis initiated in the embryo?;TIG 12 (6),1996.

46. David A.Sasson: Myogenic Regulatory Factors: Dissecting their role and regulation during vertebrate embryogenesis; Developmental Biology 156, 11-23,1993.

47. Yuki Sukiyama, Atsushi Suzuki, Masaru Kishikawa, Rika Akutsu, Tomonori Hirose, Mary M. Y. waye, Stephan K.W. Tsui, Shosei Yoshida and Shigeo Ohno: Muscle develops a specific form of small heat shock protein complex composed of MKBP/HSPB2 and HSPB3 during myogenic differentiation; The journal of Biological chemistry 275 (2),1095-1104,2000.

48. Shyamal K. Goswami, Ana Maria Zarraga, Mary Ellen Martin, Donald Morgenstern and M.A.Q.Siddiqui: Fos-mediated repression of cardiac myocyte light chain-2 gene transcription; Cellular and Molecular biology 38 (1), 049-058, 1992.

49. . M.A.Q.Siddiqui and Shyamal K Goswami: Molecular basis of Cardiogenesis; The developing heart, 9-22, 1997.

50. Eric B. Vincent, Raymond B. Runayn and Daniel L.Weeks: Production of the transforming growth factor-b binding protein Endoglin is regulated during chick heart development; Development Dynamics 213, 237-247 1998.

51. Wade R, Kedes L: Developmental regulation of contractile protein genes; Annu. Rev. Physiol. 51, 179-188, 1989.

52. Gary E Lyons: Vertebrate development; Curr. Opin. Genet. and Development, 6: 454-460, 1996.

53. Fishman, M.C. &D.Y.Staine: Cardiovascular development. Prospects for a genetic approach; Cir. Res. 74, 757-763, 1994.

54. Katherine E.Yutzey, David Barder: Diversification of cardiomyogenic cell linagees during early heart development; Circ. Res 77, 216-219, 1995.

55. Bodmer R, and Venkatesh T V Heart development in drosophila and vertebrates: conservation of molecular mechanisms; Dev. Genet. 22 (3) 181-186, 1998.

56. Nadal - Ginard B and Mahdavi V: Basic mechanism of cardiac gene expression:;European Heart Journal 14, 2-11, 1993.

57. Bernard Swynghedauw: Developmental and functional adaptation of contractile proteins in cardiac and skeletal muscles; Physiological Reviews 66 (3), 1986.

58. Vittorio Sartorelli, Masahiko Kurabayashi, and larry Kedes: Muscle-specific gene expression ; A comparison of cardiac and skeletal muscle transcription strategies; Circulation Research 72 (5), 925-931, 1993.

59. Carol A. Eisenberg and Leonard M.Eisenberg: WNT11 promotes cardiac formation of early mesoderm; Developmental Dynamics 216, 45-58,1999.

60. Makoto Tanaka, Zhi Chen, Sonia Bartunkova, Naohito Yamaski and Seigo Izumo: The cardiac homeobox gene Csx/Nkx 2.5 lies genetically upstream of multiple genes essential for heart development; Development 126. 1269-1280 1999.

61. Layons I, Parsons L M, Hartley L, Li R, Andrews J E, Robb L and Harvey RP: Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the Homeobox gene Nkx2.5; Genes and Dev. 9 (13), 1654-1666, 1995.

62. Tary J, Lints, Linda M. Parsons, Lynne Hartley, Ian Lyons and Richard P. Harvey:NKx2.5: A novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants ; Development 19 ,419-431,1993.

63. Markoto Tanka, Stephanie Burns Wechsier, Ike W. Lee, Naohito Yamaski, joel A, Lawitts and Seigo Izumo: Complex modular cis-acting elements regulate expression

of the cardiac specifying homeobox gene Csx/Nkx2.5; Development 126, 1439-1450 ,1999.

64. Ching-Ling Lien, Chunzhen Wu, Brian Mercer, Robert Webb, James A.Richardson and Eric N.Olson: Control of early-specific transcription of Nkx2.5 by a GATA-dependent enhancer; Development 126, 75-84 1999.

65. Eric C.Svensson, Rachel L.Tufts, Christine E.Polk and Jeffery M.Leiden: Molecular cloning of FOG-2; A modulator of transcription factor GATA-4 in cardiomyocytes; Proc. Natl. Acad. Sci. USA, 96, 959-961, 1999.

66. Sergei G.Tevosian, Anne E.Deconinck, Markoto Tanaka, Martina Schinke, Silvio H. Litovsky, Seigo Izumo, Yuko Fujiwara and Stuart H. Orkin: FOG-2 a cofactor for GATA transcription factors, is essential for heart morphogenesis and development of coronary vessels form epicardium; Cell 101, 729-739, 2000.

67. Tatsuya Morimoto, Koji Hasegawa, Satoshi Kaburagi, Tsuyoshi Kakita, Hiroshi Masutani, Richard N.Kitsis, Akira Matsumori and Shigetake Sasayama: GATA-5 is involved in Leukemia inhibitory factor- responsive Transcription of the beta-myosin chain gene in cardiac Myocytes; The Journal of Biological Chemistry 274 (18) , 12811- 12818, 1999.

68. Manoussos Koutsourakis, An Langeveld, Roger Patient, Rosa Beddington and Frank Grosveld: The transcription factor GATA6 is essential for early extraembryonic development; Development 126, 723-732, 1999.

69. Shrivastava D: HAND Proeins : molecular mediators of cardiac development and congenital heart disease; Trends Cardoivasc. Med, 9 (1-2), 11-18, 1999.

70. Maha Abdellatif and Michael D. Schneidert: An Effector-like function of Ras GTPase-activating Protein predominates in Cardiac Muscle Cells; The Journal of Biological Chemistry 272 (1), 525-533, 1997.

71. Guoqing Chen, Albert J.Courey: Groucho/TLE family proteins and transcriptional repression; Gene 249,1-16, 2000.

72. Shyamal K Goswami and M.A.Q Siddiqui: Molecular basis of cardiocyte cell specification; Myocardial Preservation, Preconditioning and Adaptation 293 259-266, 1996.

73. Marko E. Horb and Gerald H.: Thomsen.Tbx5 is essential for heart development; Development 126, 1739-1751,1999.

74. Olson and Srivastava D: Molecular pathways controlling heart development; Science 272 (5262) 617-676, 1996.

75. Ptterson K.D, Cleaver O, Gerber W V, Grow M W, Newman C.S and Krieg P.A: Homeobox gene in cardiovascular development; Current Opinion in Developmental Biology 40, 1-44, 1998.

76. Subramanian S V and Nadal-Ginard B: Early expression of the different isoforms of the myocyte enhancer factor-2 (MEF-2) protein in myigenic as well as non-myogenic cell lineages during mouse embryogenesis; Mech. Dev. 57 (1) 103-112, 1996.

77. Snyder JM, Mendeleson CR, Johnston JM.: The morphology of lung development in the human fetus. Pulmonary development: Transition from intrauterine to extrauterine life, ed GH Nelson, pp. 19-46, 1985.

78. Shanon JM: Induction of alveolar type II cell differentiation in fetal tracheal epithelium by grafting distal lung mesenchyme; Dev. Biol. 166, 600-14, 1994.

79. Shanon JM, Nielson LD, Gebb SA, Randell SH: Mesenchyme specifies epithelial differentiation in receprocal recombination of embryonic lung and trachea; Dev. Dyn. 212, 482-494, 1998.

80. Bellusci S, Grindley J, Emoto H, Itoh N, Hogan BLM: Fibroblast growth factor 10 (FGF 10) and branching morphogenesis in the embryonic mouse lung; Development 124, 4867-78, 1997.

81. Park WY, Miranda B, Lebeche D, Hashimoto G, Cardoso WV: FGF-10 is a chemotactic factor for distal epithelial buds during lung development;Dev. Biol. 201, 125-34, 1998.

82. Grindley JC, Bellusci S, Perkins D, Hogan BL: Evidence for the involvement of the Gli gene family in embryonic development; Dev. Biol. 188, 337-348, 1997.

83. Motoyama J, Liu J, Mo R, Ding Q, Post M, Hui CC: Essential function of Gli2 and Gli3 in the formation of lung, trachea, and oesophagus; Nat. Genet. 20, 54-57, 1998.

84. Dolle P, Ruberte E, Leroy P, Morriss-Kay G, Chambon P: Retinoic acid receptors and cellular retinoid binding proteins. I. Asystematic study of their differential pattern of transcription during mouse organogenesis; Development 110, 1133-51, 1990.

85. Sukov HM, Izpisua-Belmonte JC, Ganan Y, Evans RM: Mouse embryo lacking RXR alpha are resistant to retinoic-acid-induced limb defects; Development 121, 3997-4003, 1996.

86. Kastner P, Grondona JM, Mark M, Gansmuller A, LeMeur M: Genetic analysis of RXR alpha developmental function: convergence of RXR and RAR signaling pathways in heart and eye morphogenesis; Cell 78, 987-1003, 1994.

87. Massague J: TGF-beta signal transduction; Annu. Rev. Biochem. 67, 753-91, 1998.

88. Stanton BR, Perkins AS, Tessarollo L, Sassoon DA, Parada LF: Loss of N-myc functions results in embryonic lethality and failure of the epithelial component of the embryo to develop; Genes Dev. 6, 2235-47, 1992.

89. Mones CB, Auerbach AB, Conlon RA, Joyner AL, Rossant J: A targeted mutation reveals a role for N-myc in branching morphogenesis in the embryonic mouse lung; Genes Dev. 6, 691-704, 1992.

90. Blackwood EM, Eisenman PN: Max: a helix-loop-helix zipper protein that forms sequence-specific DNA binding complex with Myc; Science 251, 1211-17, 1991.

91. Lazzaro D, Price M, De Felice M, Di Lauro R: The transcription factor TTF-1 is expressed at the onset of thyroid and lung morphogenesis and in restricted regions of fetal brain; Development 113, 1093-104, 1991.

92. Kikkawa F, Gonzalez FJ, Kimura S: Characterization of a thyroid-specific enhancer located 5.5 kilobase pairs upstream of human thyroid peroxidase gene; Mol. Cell. Biol. 10, 6216-24, 1990.

93. Civitareale D, Lonigro R, Sinclair AJ, Di, Lauro R: A thyroid-specific nuclear protein essential for tissue expression of the thyroglobulin promoter; EMBO J. 8, 2537-42, 1989.

94. Carole R. Mendelson: Role of transcription factors in fetal lung development and surfactant protein gene expression; Annu. Rev. Physiol. 62, 875-915, 2000.

95. McKenna NJ, Lanz RB, O'Melly BW: Nuclear receptor coregulators; cellular and molecular biology Endo. Rev. 20, 321-44,1995.

96. cole TJ, Blendy JA, Monaghan P, Kriegestein K, Schmid W: Targeted disruption of the glucocorticoid receptor gene blocks adrenergic chromaffin cell development and severely retards lung maturation; Gene Dev. 9, 1608-21, 1995.

97. Reichardt HM, Kaestner KH, Tucher-mann J, Kretz O, Block R, Gass P, Schmid W, Herrlich P, Angel P, Schutz G: DNA binding of the glucocorticoid receptor is not essential for survival; Cell 93, 531-41, 1998.

98. Muglia L, Bae DS, Brown TT, Vogt SK, Alvarez JG: Proliferation and differentiation defects during lung development in corticotropin-releasing hormone-deficient mice; Am. J. Respir. Cell. Mol. Biol. 20, 181-88, 1999.

99. Muglia L, Jacobson L, Dikkes P, Majoub JA: Corticotropin-releaseing hormone defeciency reveals major fetal development but not adult glucocorticoid need; Nature 373 (6513), 427-32, 1995.

100. Odom MJ, Synder JM, Boggaram V, Mendelson CR: Glucocorticoid regulation of the major surfactant-associated protein (SP-A) and its mRNA and of morphologic development of human fetal lung in vitro; endocrinology 123, 1712-20, 1988.

101. Brindle PK, Montminy MR: The CREB family of transcription activators; Curr. Opin. Genet. Dev. 2, 199-204, 1992.

102. Rodoloph D, Tafuri A, Glass P, Hammer-ling GJ, Arnold B, Schultz G: Impaired fetal T cell development and perinatal lethality in mice lacking the cAMP response element binding protein; Proc. Natl. Acad. USA 95, 4481-86,1998.

103. Botas C, Poulain F, Akiyama J, Brown C, Allen L: Alterd surfactant homeostasis and alveolar type II cell morphology in mice lacking surfactant protein D; Proc. Natl.Acad. Sci. USA 95, 11869-74, 1998.

104. Korfhagen TR, Sheftelyevich V, Burhans MS, Bruno MD, Ross GF: Surfactant protein-D regulates surfactant phospholipid homeostasis in vivo; J. Biol. Chem. 273, 28438-43, 1998.

105. Kan Zaret:Early liver differenciation: genetic potrentiation and multilevel growth control:CurrentOpinion in genetics and development 8:526-531,1998.

106. Le Douarin NM: An experimental analysis of liver development; Med Biol 53:427-455.,1975.

107. Gualdi R,Bossard P, Zheng M, Hamada Y, Coleman JR,Zaret KS: Hepatic specification of the gut endoderm in vitro: cell signalling and transcriptional control;GenesDev 10:1670-1682,1996.

108. Cascio S, Zaret KS: Hepatocyte differenciation initiates during endodermal-mesenchymal interactions prior to liver formation;Development,113:217-225,1991.

109. Shiojiri N,Lemire JM, Fausto N: Cell lineages and oval cell progenitors in rat liver development.Cancer Res; 51: 2611-2620,1991.

110. Shiojiri N: Analysis of differentiation of hepatocytes and bile duct cells in developing mouse liver by albumin immunoflorescence ; Dev Growth Differ, 26 :555-561,1984.

111. Weinstein DC, Ruiz i Altaba A, Chen WS, Hoodless P, Prezioso VR.Jessel TM, Darnell JE Jr: The winged- helix transcription factor HNF-3 beta is required for notochord development in the mouse embryo; Cell 78 :575-588,1994.

112. Kuo CT,Morrisey EE , Anandappa r, sigrist K, Lu MM, Parmacek MS, Soudais C,Leiden JM: GATA-4 transcription factor is required for ventral morphogenesis and heart tube formation ;Genes Dev,11:1048-1060,1997.

113. Molkentin JD, Lin Q, Duncan SA , Olson EN: Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis; Genes DEv 11 1061-1072,1997.

114. Clark KL, Halay ED , Lai E , Burley SK: Co-crystal structure of the HNF3/fork head DNA recognition motif resembles histone H5; Nature 364, 412-420,1993.

115. Zhu J , Hill RJ , Heid PJ FukuyamaM ,Sugimoto A , Priess JR , Rothman JH: *End-1* encodes an apparent GATA factor that specifies the endoderm precurser in *Ceanorhabditis elegans* embryos; Genes Dev,11:2883-2896.,1997.

116. Hilberg F, Aguzzi A , Howells N, Wagner EF: c-jun is essential for normal mouse development and hepatogenesis; Nature 365;179-181,1993.

117. Hentsch b, Lyons I , Ruili L , Hartly L, Lints TJ , Adams JM, Harvey RP: Hlx homeo box gene is essential for an indusive tissue interaction that drives expansion of embryonic liver and gut; Genes Dev 10:70-79,1996.

118. Montesano R, Matsumoto K, Nakamura T , Orci L : Identification of a fibroblast- derived epithelial morphogen as hepatocyte growth factor; Cell 67, 901-908, 1991.

119. Sukov HM , Dyson E , Gumeringer CL, Price J , Chien KR, Evans RM : RXR alpha mutant mice establish a genetic basis for vitamin a signaling in heart morphogenesis.Genes Dev 8 :1007-1018,1994.

120. Girish N.Nallur,Kulkarni Prakash,and Sherman M Weissman: Multiplex selection technique(MuST):an approach to clone transcription factor binding sites; Proc.Natl.Acad.Sci,USA, 93, -1184-1189, 1996.