

REGULATION OF RAT SMAD1 GENE EXPRESSION

BY

Yu Sun

A Thesis submitted to

The Faculty of Graduate Studies

In Partial Fulfillment of the Requirements for the Degree of Master of Science

Faculty of Pharmacy

University of Manitoba

Winnipeg, Manitoba

© Yu Sun, July 2006

THE UNIVERSITY OF MANITOBA
FACULTY OF GRADUATE STUDIES

COPYRIGHT PERMISSION

REGULATION OF RAT SMADI GENE EXPRESSION

BY

Yu Sun

**A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of
Manitoba in partial fulfillment of the requirement of the degree
OF**

MASTER OF SCIENCE

Yu Sun © 2006

Permission has been granted to the Library of the University of Manitoba to lend or sell copies of this thesis/practicum, to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film, and to University Microfilms Inc. to publish an abstract of this thesis/practicum.

This reproduction or copy of this thesis has been made available by authority of the copyright owner solely for the purpose of private study and research, and may only be reproduced and copied as permitted by copyright laws or with express written authorization from the copyright owner.

ABSTRACT

Introduction:

Hepatic stellate cell activation is the central event of liver fibrosis. Our previous studies revealed the potential role of Smad1 in hepatic stellate cell activation. Smad1 is an intracellular protein mediating signal transduction of transforming growth factor family. In the present study, we cloned rat Smad1 5'-flanking region and intron1 region and investigated the transcriptional activity.

Methods:

Polymerase Chain Reaction was used to clone rat Smad1 fragments. Dual Luciferase Assay was performed to study the transcriptional activity. Electrophoretic Mobility Shift Assay was used to determine the DNA/protein interaction.

Results:

Rat Smad1 had two promoters at its 5'-flanking region and intron1 region respectively. 5'-flanking promoter was CG-rich containing two Sp1 binding sites. Transcriptional activity of 5'-flanking promoter was higher than intron1 promoter in Chang liver cells and CFSC-8B cells. The activities of both promoter regions were greater in activated than in quiescent CFSC-8B cells. Several positive and negative regulatory elements were identified in both regions. Oct-1 binding site was defined in intron1 promoter.

Conclusions:

Rat Smad1 contained a CG-rich promoter in its 5'-flanking region and a second promoter in its intron1 region. Both promoters were highly active in activated CFSC-8B cells. Binding site of Oct-1 factor was identified in intron1 promoter.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Yuewen Gong, for providing such an interesting project. I had little background of molecular biology when I first started. I really appreciate your patience in encouragement of my study. You really taught me how to do research correctly. Especially, I would like to express my gratitude for your hard work on my thesis.

Thanks to my committee members Dr. Mike Namaka and Dr. Gerald Minuk for your great suggestions on my research project. I always remember the time when Dr. Mike Namaka asked me about the leading group in my research field. That was the moment I realized there was so much more I should know other than myself. It was a great honor and joy to work with Dr. Gerald Minuk. I could never forget all the inspired questions you challenged me, which also let me know how intelligent you are. I really appreciate all the work you have done for me.

I would like to thank my lab mates, Ms. Jianghong Fan, Dr. Hong Shen, Ms. Jingfen Qi, Ms. Jing Yan and Ms. Yu Gong. You guys made our lab like a big family. To meet you in the lab was the happiest thing for my day. I have been through so much with Ms. Fan and I always believe you are the friend for life.

At last, I would like to thank my parents for your understanding and supports. Although you are far away in China, I can always feel you love with me!

TABLE OF CONTENTS

ABSTRACT	i
ACKNOWLEDGEMENTS	ii
TABLE OF CONTENTS	iii
LIST OF TABLES	vi
LIST OF FIGURES.....	vii
LIST OF ABBREVIATIONS	x
 I. INTRODUCTION	 1
1. Liver Fibrosis	1
1.1 General aspects	1
1.2 Pathogenesis	2
2. Hepatic Stellate Cells.....	5
2.1 History.....	5
2.2 Quiescent HSCs in normal liver	6
2.3 Activated HSCs in diseased liver	7
2.4 Factors contribute to activation of HSCs	7
3. Transforming Growth Factor- β	10
3.1 Ligand and receptor	10
3.2 TGF- β activation	11
3.3 TGF- β 1 on liver fibrosis	12
4. Smad Family	15
4.1 History.....	15
4.2 Subtype.....	15
4.3 Structure.....	16
4.4 Signaling pathway	18
5. Smad1 Gene	22

5.1	<i>Comparison of Smad1 gene</i>	22
5.2	<i>Rat Smad1 gene</i>	24
6.	Gene Expression Regulation	28
6.1	<i>Regulation of gene expression</i>	28
6.2	<i>Control of transcriptional activity</i>	28
6.3	<i>Types of promoters</i>	31
7.	Previous Studies	33
7.1	<i>Smads and liver fibrosis</i>	33
7.2	<i>Smad1 and liver fibrosis</i>	34
II.	MATERIALS AND METHODS	37
1.	Materials	37
2.	Methods	38
2.1	<i>Genomic DNA Isolation</i>	38
2.2	<i>Restriction Enzyme Digestion</i>	39
2.3	<i>Agarose Gel Electrophoresis</i>	39
2.4	<i>PCR Cloning</i>	40
2.5	<i>Reconstruction of plasmid</i>	41
2.6	<i>Bacterial Transformation</i>	42
2.7	<i>Plasmid DNA Isolation</i>	43
2.8	<i>Cell Culture and morphology</i>	44
2.9	<i>Transient Transfection</i>	46
2.10	<i>Protein Isolation</i>	46
2.11	<i>Dual Luciferase Assay</i>	47
2.12	<i>Western blot analysis</i>	49
2.13	<i>DNA Labeling and Purification</i>	50
2.14	<i>Electrophoretic Mobility Shift Assay (EMSA)</i>	51
2.15	<i>Sequence Analysis</i>	51
III.	RESULTS	53
1.	Rat genomic DNA isolation	53

2. Cloning of rat Smad1 5'-flanking region and intron1 region.....	54
2.1 PCR cloning.....	54
2.2 Subcloning and reconstruction of plasmid.....	59
3. Transcriptional activity Analyses of rat Smad1 gene.....	74
3.1 Identification of rat Smad1 promoter regions.....	76
3.2 Deletional analyses of rat Smad1 5'-flanking region and intron1 region....	78
3.3 Comparison of the activities in Chang liver cells and CFSC-8B cells.....	80
3.4 Comparison of the activities in activated and quiescent CFSC-8B cells.....	82
3.4.1 Morphology and viability of CFSC-8B cells cultured on Matrigel™.....	82
3.4.2 Transcriptional activity of rat Smad1 in quiescent CFSC-8B cells.....	85
4. Investigation of transcriptional factors in rat Smad1 5'-flanking region.....	86
4.1 Characterization of rat Smad1 5'-flanking region.....	86
4.2 Interactions between trans-factors and cis-elements in 5'-flanking region.....	89
5. Investigation of transcriptional factors in rat Smad1 intron1 region.....	93
5.1 Characterization of rat Smad1 intron1 region.....	93
5.2 Interactions between trans-factors and cis-elements in intron1 region.....	99
5.3 Determination of trans-factors in intron1 region.....	100
IV. DISCUSSIONS.....	103
1. 5'-Flanking Promoter Region.....	104
1.1 CG-rich promoter.....	104
1.2 C/EBP element.....	106
2. Intron1 Region.....	106
2.1 Two transcripts of rat Smad1.....	106
2.2 Transcriptional activity of the intron1 region.....	107
2.3 Oct-1 binding site.....	108
3. Transcriptional activity of Smad1 gene in HSCs.....	109
4. Smad transcriptional regulation.....	109
V. REFERENCES.....	112

LIST OF TABLES

Table1. Materials.....	37
Table2. Cell culture conditions.....	45
Table3. Websites of sequence analysis.....	52
Table4. Primer sequences and PCR conditions of cloning of rat Smad1 5'-flanking and intron1 region.....	55
Table5. Potential transcriptional factors for F-430/+1 sequence.....	88
Table6. Information of EMSA probes.....	91
Table7. PCR conditions of EMSA probes.....	91
Table8. Potential transcriptional factors for I-273/+1 sequence.....	95
Table9. Potential transcriptional factors for sequences of rat Smad1 intron1 region.....	98
Table10. Consensus sequences of putative transcription factors in Intron1 region.....	101

LIST OF FIGURES

Fig.1. Hepatic architecture.....	3
Fig.2. Activation of HSCs.....	8
Fig.3. The effects of TGF- β on fibrogenesis.....	14
Fig.4. Schematic structures of Smads.....	17
Fig.5. Signaling pathway of TGF- β family.....	20
Fig.6. Receptors and intracellular mediators of TGF- β family.....	21
Fig.7. Smad1 gene in different species.....	23
Fig.8. Alignment analysis of two rat Smad1 transcripts.....	25
Fig.9. An overview of rat Chromosome 19.....	26
Fig.10. Rat Smad1 gene information from Ensembl website.....	27
Fig.11. 5'-flanking region of a gene.....	29
Fig.12. Comparison of three types of eukaryotic RNA polymerase II promoters..	33
Fig.13. Bioluminescent reactions catalyzed by firefly and <i>Renilla</i> luciferases.....	48
Fig.14. Rat genomic DNA analysis.....	53
Fig.15. Schematic organization of cloning of rat Smad1 5'-flanking region and intron1 region.....	56
Fig.16. PCR cloning of rat Smad1 5'-flanking region.....	58
Fig.17. PCR cloning of rat Smad1 intron1 region.....	58
Fig.18. pGL2-Basic vector map.....	60
Fig.19. pCP4-Topo vector map.....	61
Fig.20. Smad1.pGL2 constructs containing fragments of rat Smad1 5'-flanking region	62

Fig.21. Smad1.pCR4-Topo constructs containing fragments of rat Smad1 intron1 region.....	63
Fig.22. Smad1.pGL2 constructs containing fragments of rat Smad1 intron1 region	63
Fig.23. Sequences alignment.....	73
Fig.24. CollagenX.pGL2 map.....	74
Fig.25. pRL-TK vector map.....	75
Fig.26. Transcriptional activity analyses of rat Smad1 5'-flanking region and intron1 region in Chang liver cells.....	79
Fig.27. Deletional analyses of rat Smad1 5'-flanking region and intron1 region	79
Fig.28. Comparison of activities in Chang liver cells and CFSC-8B cells.....	81
Fig.29. Morphology of CFSC-8B cells cultured on plastic dish and Matrigel™ dish.	83
Fig.30. Expression of α -Smooth muscle actin (α -SMA).....	84
Fig.31. Transcriptional activities of rat Smad1 in quiescent CFSC-8B cells.....	85
Fig.32. Sequence analysis of F-430/+1 of rat Smad1 5'-flanking region.....	87
Fig.33. Schematic organization of rat Smad1 probes for EMSA.....	90
Fig.34. PCR cloning of EMSA probe2.....	92
Fig.35. The interaction between nuclear proteins and Probes of rat Smad1 5'-flanking region.....	92
Fig.36. Sequence analysis of I-273/+1 of rat Smad1 intron1 region.....	94
Fig.37. Sequence analysis of I-389/-273 of rat Smad1 intron1 region.....	96

Fig.38. Sequence analysis of I-610/-490 of rat Smad1 intron1 region.....	96
Fig.39. Sequence analysis of I-490/-389 of rat Smad1 intron1 region.....	97
Fig.40. PCR cloning of EMSA probe3, Probe4 and Probe5.....	99
Fig.41. The interaction between nuclear proteins and Probes of rat Smad1 intron1 region.....	100
Fig.42. Competition assay using GATA, E2F and Oct-1 consensus sequences	102

LIST OF ABBREVIATIONS

°C	degrees Celsius
α -SMA	α -Smooth muscle actin
BSA	bovine serum albumin
CO ₂	carbon dioxide
DTT	Dithiothreitol
DNA	deoxyribonucleic acid
EB	ethidium bromide
<i>E.coli</i>	<i>Escherichia coli</i>
ECL	Electrogenerated chemiluminescence
EDTA	ethylene diaminetetracetic acid
EMSA	Electrophoresis Mobility Shift Assay
FBS	fetal bovine serum
g	Gram
HEPES	hydroxyethylpiperazine ethanesulfonic acid
HSC	hepatic stellate cell
IgG	Immunoglobulin G
kDa	kilo Daltons
LB	Luria-Bertani
M	Molar
mM	Millimolar
μ g	Microgram
min	minute(s)
ml	Milliliter
μ l	Microliter
NaCl	sodium chloride
NaOAc	sodium acetate
NaOH	sodium hydroxide
OD	optical density
O/N	Overnight
PBS	phosphate buffered saline
PCR	Polymerase Chain Reaction
rpm	rotations per minute
RNA	ribonucleic acid
SDS	sodium dodecyl sulfate
TBS	Tris buffered saline
TBS-T	Tris buffered saline-Tween20
TE	Tris-EDTA
v	Voltage
UV	Ultraviolet

I. INTRODUCTION

1. Liver Fibrosis

1.1 General aspects

Liver fibrosis is a common sequel to different chronic liver disease (Albanis and Friedman 2001), such as viral hepatitis, autoimmune hepatitis, drug-induced liver disease, alcoholic or non-alcoholic steatohepatitis (NASH) and genetic liver diseases. Of these, the most common causes of liver fibrosis and cirrhosis are chronic HCV infection, alcohol abuse, and NASH (Friedman 2003; Brunt 2004).

The onset of liver fibrosis is usually insidious and the most of related morbidity and mortality appear after the development of cirrhosis (Poynard, Ratziu et al. 2000). Cirrhosis, characterized by nodule formation and altered hepatic function (Friedman 2003), is defined as the end stage of fibrosis by Metavir scoring system, which assesses progression of liver fibrosis (Poynard, Bedossa et al. 1997). Most patients have an interval of 15-20 years to progress to cirrhosis but several clinical situations can accelerate this process, such as repeated acute alcoholic hepatitis, subfulminant hepatitis, and fibrosing cholestasis in patients with HCV re-infection after liver transplantation (Bataller and Brenner 2005). The clinical manifestations of cirrhosis are various from no symptoms at all to liver failure. Major complications of cirrhosis include ascites, variceal hemorrhage, renal failure, hepatic encephalopathy (Gines, Cardenas et al. 2004). Patients with cirrhosis can remain asymptomatic for several years, but progressive deterioration is inevitable once complications develop (Davis, Albright et al. 2003). In addition, cirrhosis is a risk factor for developing hepatocellular carcinoma (Sarbah, Gramlich et al. 2004).

Over the past three decades, our knowledge of hepatic fibrosis has been increased (Bataller and Brenner 2005). The growing awareness of molecular

underpinnings of fibrosis provides not only new theories of pathogenesis, but also methods of detection and treatment of the disease. However, at present, the best method for assessing fibrosis is liver biopsy (Afdhal and Nunes 2004) so called the 'gold standard'. The most effective therapy for liver fibrosis is to remove the causative agent (Bataller and Brenner 2001; Bataller and Brenner 2005). Liver fibrosis was historically thought to be a passive and irreversible process (Popper and Uenfriend 1970). However, ample evidence now demonstrates that the reversal of liver fibrosis or even cirrhosis, is often possible (Soyer, Ceballos et al. 1976; Hammel, Couvelard et al. 2001).

1.2 Pathogenesis

Regardless of etiology, liver fibrosis is characterized by disruption of hepatic architecture and depositing of excess abnormal extracellular matrix (ECM) constituents, including fibrillar collagens, sulfated proteoglycans and glycoproteins (Rojkind, Giambrone et al. 1979).

The liver lobule is formed by parenchymal cells and nonparenchymal cells (Sasse, Spornitz et al. 1992). Hepatocytes (one of the hepatic parenchymal cells) made up over 80% of the volume of the human liver and perform the majority of numerous liver functions. Normal hepatocytes are cuboidal in shape and arranged in palisades along the sinusoid. Nonparenchymal liver cells, which contribute only 6.5% to the liver volume, but 40% to the liver cells number, lie in the sinusoidal compartment of the tissue (Kmieć 2001).

The hepatic sinusoid is formed by several cell types (Fig.1.A). The vascular sinusoid is lined by highly specialized sinusoid endothelial cells (SECs). Sinusoid endotheliums in the liver consist of numerous fenestrae, which allow free passage of

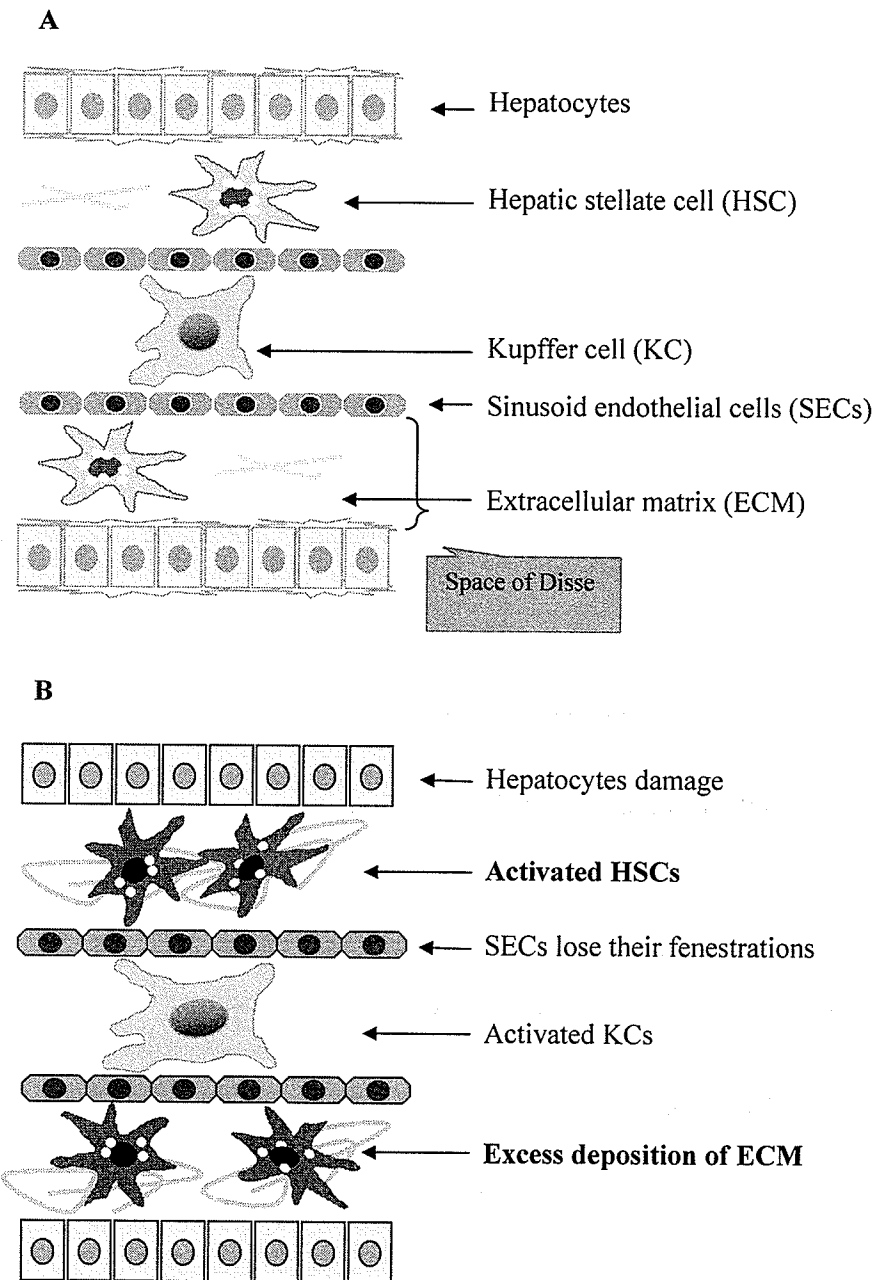


Fig.1. Hepatic architecture. A. Normal liver. B. Diseased liver. Following chronic liver injury, some of the hepatocytes are necrotic or apoptotic. Kupffer cells activate. Hepatic stellate cells undergo phenotypical change featuring secretion of excess extracellular matrix proteins. Sinusoid endotheliums lose their fenestrations.

macromolecules from sinusoidal blood into the space of Disse, thus permitting contact with hepatocytes (Arii and Imamura 2000). In the sinusoidal lumen, there are Kupffer cells (KCs), which are the resident tissue macrophages responsible for removing bacterial antigens and endotoxin (Catala, Anton et al. 1999). Between the sinusoidal endotheliums and the sinusoidal surface of hepatocytes lies the space of Disse, containing hepatic stellate cells (HSCs) and basement membrane-like extracellular matrix (Bouwens, De Bleser et al. 1992).

Two principal types of extracellular matrix (ECM) proteins are identified in the liver: fibrillar or interstitial matrix proteins and basement membrane-like matrix proteins (Neubauer, Saile et al. 2001). Fibrillar matrix proteins including collagen I, III and V are located in the liver capsule, the large and medium sized blood vessels and the portal tracts. Basement membrane-like matrix proteins including collagen IV, laminin and proteoglycans lie in the functionally important sinusoidal space of Disse. Type IV collagen and laminin and proteoglycans form a flexible lattice structure that provides cellular support and allows unimpeded transport of solutes and growth factors. In addition basement membrane-like matrix proteins provide signals that maintain the differentiated function of surrounding cells. The degradation of ECM is controlled by matrix metalloproteinases (MMPs), a family of zinc and calcium dependent endopeptidases (Murphy, Docherty et al. 1991). Tissue inhibitors of metalloproteinases (TIMPs) are the inhibitors of active MMPs, acting as an important regulatory brake on MMPs activity (Docherty, Lyons et al. 1985).

Liver fibrosis is currently considered as a model of the wound-healing response to repeated liver injury (Albanis and Friedman 2001). After an injury, hepatocytes regenerate and replace the necrotic and apoptotic cells. This process is associated with an inflammatory response and a limited ECM deposition (Bataller and Brenner

2005). However, when the hepatic injury persists (Fig.1.B), the liver regeneration eventually fails. Damaged hepatocytes and infiltrated inflammatory lymphocytes activate Kupffer cells, which in turn release numerous cytokines regulating the behavior of surrounding cells (Wheeler 2003). Sinusoid endothelial cells lose their fenestrations and hepatic stellate cells undergo a phenotypical activation and differentiate into myofibroblast-like cells (Kinnman and Housset 2002). Activated HSCs migrate and accumulate at the sites of tissue injury, where they secrete large amounts of ECM proteins, including collagen I, III and IV (Burt, Griffiths et al. 1990), laminin, fibronectin (Schuppan 1990), undulin, elastin (Gressner and Bachem 1990), hyaluronan and proteoglycans (Kovalszky, Nagy et al. 1998). In advanced fibrotic stages, the liver contains approximately six times the amount of matrix as normal liver (Schuppan 1990). Accumulation of ECM is resulted from both increased synthesis and decreased degradation. In the space of Disse, the amounts of collagen I, III and IV are all increased, but the increase of collagen I is disproportionate, so the ratio of collagen I to collagen III and IV augments (Alcolado, Arthur et al. 1997). Thus, the ECM composition becomes scar-like. The appearance of hepatic sinusoid alters and hepatocellular function deteriorates.

Liver fibrosis is a complex process, involving in numbers of cell types and numerous factors, but central event among these has been identified as the activation of hepatic stellate cells (Albanis and Friedman 2001).

2. Hepatic Stellate Cells

2.1 History

In 1876, von Kupffer described hepatic Sternzellen, which was a type of star-shaped cells in the liver, but the function of these cells remained enigmatic for

decades. In 1951, Ito observed lipid-containing perisinusoidal cells in human liver. In 1971, Wake demonstrated that the Sternzellen described by von Kupffer and the lipid-containing cells described by Ito were identical and established that these cells were important sites of vitamin A storage (Geerts 2001). Hepatic stellate cells received little attention until the 1980s, when Friedman et al presented that these cells were major collagen-producing cells in the liver and intimately linked to the pathogenesis of hepatic fibrosis (Friedman, Roll et al. 1985). Since then, HSCs have been studied in detail and become the central cell type involved in liver fibrosis.

2.2 Quiescent HSCs in normal liver

HSCs, formerly termed as lipocytes, Ito cells or perisinusoidal cells are located between parenchymal cell plates and sinusoidal endothelial cells in the space of Disse, remain quiescent in normal liver.

In physiological conditions, HSCs play a cardinal role in retinoid storage and release, production and regulation of ECM molecules, and growth factor secretion (Wells 2005). Quiescent HSCs function as major retinoid-storing cells containing over 80% of total retinoid in the body and can be distinguished by prominent intracellular lipid droplets under microscope (Hendriks, Blaner et al. 1988). HSCs control extracellular matrix turnover by secreting correct amounts of a limited number of ECM molecules, and by releasing matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs) (Sarem, Znaidak et al. 2006). They are also important sources of numbers of factors that maintain homeostasis in the microenvironment of the hepatic sinusoid. In addition, HSCs are presumably related to the control of blood flow through the sinusoidal capillaries (Friedman 1990) due to their long cytoplasmic processes with fine branching.

2.3 Activated HSCs in diseased liver

In chronic liver disease, HSCs undergo transdifferentiation from quiescent phenotype to myofibroblast-like phenotype, which express α -smooth muscle actin (Wu and Zern 2000). During this activation process (Fig.2.), HSCs shed their lipid droplets, lose the cytoplasmic long processes, and acquire contractile, proliferative, and fibrogenic properties. HSC activation is the dominant event in hepatic fibrosis because activated HSCs express excess matrix proteins and downregulate their degradation by secreting TIMPs.

Activation of HSCs is divided into two stages (Reeves and Friedman 2002). First stage is defined as 'initiation', including the rapid changes in gene expression and phenotype. Initiation results from paracrine stimulation by inflammatory cells and injured hepatocytes, and from the early alteration in ECM composition. The later stage is termed 'perpetuation', including cytokine release and enhanced growth factor expression and responsiveness, which amplify the activated phenotype. Perpetuation is resulted from autocrine and paracrine stimulation.

2.4 Factors contribute to activation of HSCs

Many factors lead to activation of HSCs, such as, interplay of different cell types, remodeling of ECM, and secretion of cytokines and growth factors (Friedman 2003).

The earliest changes in HSCs reflect paracrine stimulation by all neighboring cells, including hepatocytes, Kupffer cells, and inflammatory lymphocytes (Friedman 2000). Hepatic injury, such as viral infection, alcoholic or drug toxicity, causes damage to hepatocyte and elicits an inflammatory response in the liver. The dead bodies and membrane components of the damaged hepatocytes, metabolites of toxic agents, and infiltrating inflammatory cells activate Kupffer cells (Wu and Zern 2000).

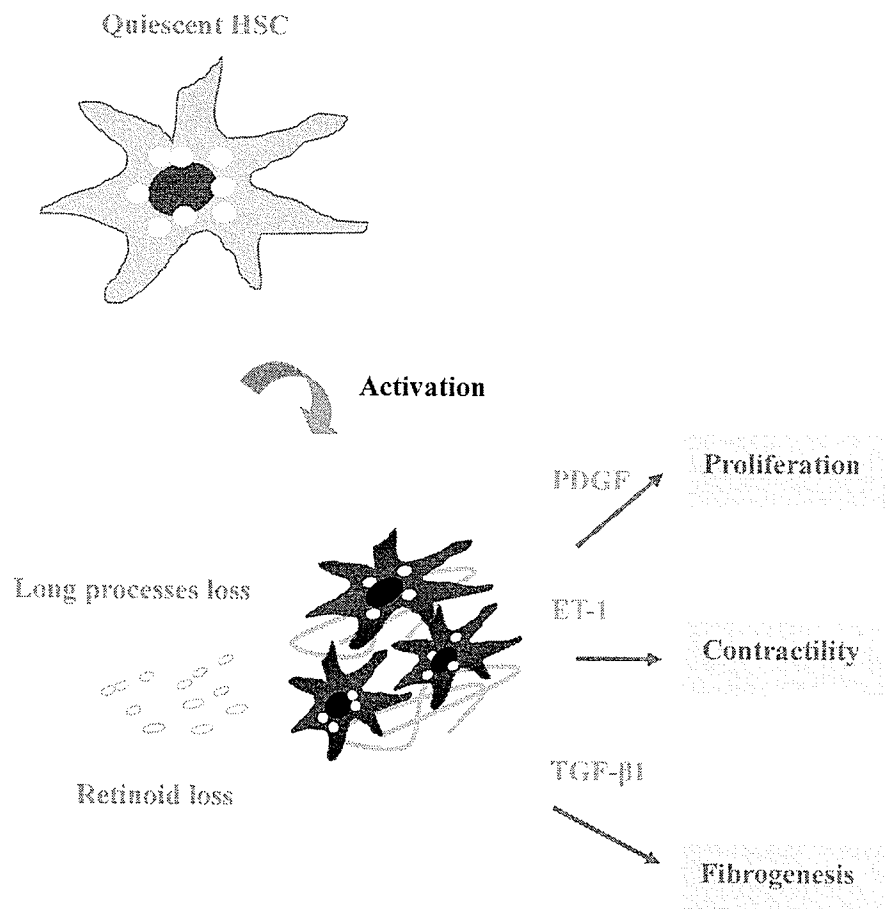


Fig.2. Activation of HSCs. During activation, HSCs lose their retinoid droplets and long processes, and acquire contractile, proliferative, and fibrogenic properties, synthesizing excess ECM components. This process involves in many cytokines and growth factors, such as, PDGF (platelet-derived growth factor), ET-1 (endothelin-1), TGF- β 1 (transforming growth factor- β 1).

Activated Kupffer cells release a number of cytokines including transforming growth factor- β (TGF- β), platelet-derived growth factor (PDGF), tumor necrosis factor- α (TNF- α), and other factors, which are important for HSC activation (Maher 2001). Soluble factors from both damaged hepatocytes and activated Kupffer cells can stimulate HSCs and induce the activation (Fallowfield and Iredale 2004).

One of the most intriguing factors mediating the phenotypic change of HSCs is the extracellular matrix (Wells 2005). In 1989, Friedman et al demonstrated that by culture on basement membrane-like matrix HSCs can be maintained in a normal quiescent phenotype (Friedman, Roll et al. 1989). However, HSCs isolated from normal liver usually go through a similar program of activation if they are cultured on an uncoated or a collagen I coated plastic tissue culture dish in the presence of serum (Geerts, Vrijssen et al. 1989; Friedman, Rockey et al. 1992). Recently, many researchers have documented that, activated HSCs can be reverted to quiescent stage by culturing on basement membrane-like matrix (Li, Sato et al. 1999; Sohara, Znoyko et al. 2002; Liu, Gaca et al. 2003). These findings reveal that alterations in the matrix substratum especially the substitution of collagen I for collagen IV are important for activation of HSCs. Methods to obtain HSCs have been standardized and prolonged culture of HSCs on plastic has been widely accepted as a method to activate the cells.

HSC activation is also influenced by a wide variety of cytokines and cell messengers, which can be paracrine or autocrine (Breitkopf, Haas et al. 2005). PDGF (platelet-derived growth factor), the most potent mitogenic stimulus to HSCs, causes HSCs proliferation during liver fibrosis (Pinzani 2002). During activation, there is a remarkable increase in the contractility of HSCs leading to increased portal resistance. Endothelin-1 (ET-1) is the key contractile stimulus to HSCs. The

expression of ET-1 and its ET-A receptor is increased during liver injury (Shah 2001). Fibrogenesis is the most important property of activated HSCs that is regulated by TGF- β 1 (transforming growth factor- β 1). TGF- β 1 has profound effects on matrix production (Gressner, Weiskirchen et al. 2002). A growing list of other factors mediating HSC activation has been identified, such as VEGF (vascular endothelial growth factor), FGF (fibroblast growth factor) and MCP-1 (monocyte chemotactic protein-1) (Friedman 2003). Some new function of the known cytokines has been documented, for instance PDGF is also a potent chemoattractants for activated HSCs (Tangkijvanich, Tam et al. 2001) other than its important role in proliferation.

Our knowledge of HSC biology has been growing quickly. The factors contributing to HSC activation are well studied, among which TGF- β 1 attracts the most attention.

3. Transforming Growth Factor- β

3.1 Ligand and receptor

The name of transforming growth factor- β is attributed to its ability to induce phenotypic transformation (Hsuan 1989). TGF- β 1 represents the main profibrogenic cytokine in liver fibrosis (Breitkopf, Haas et al. 2005).

Five isoforms of TGF- β have been identified, TGF- β 1 to TGF- β 5, of which only TGF- β 1, TGF- β 2 and TGF- β 3 have been detected in the liver (Jakowlew, Mead et al. 1991). Of the three types of isoforms, TGF- β 1 is the most well-characterized. Cellular sources of TGF- β 1 are various. TGF- β 1 can be derived from hepatic stellate cells, Kupffer cells, hepatocytes, sinusoidal endothelial cells and platelets (Okuyama, Shimahara et al. 2002). HSCs are the most important source of TGF- β 1 during liver fibrosis (Bissell, Wang et al. 1995). It has been documented that, the

expression of active TGF- β 1 is elevated not only in animal model of liver fibrosis, but also in fibrotic liver from patients with liver fibrosis and cirrhosis (Castilla, Prieto et al. 1991).

There are three main types of TGF- β receptors. Type I and type II receptors, important for its signal transduction, are present in both quiescent and activated HSCs (Dooley, Delvoux et al. 2000). The type III receptor, which has long been known as non-signaling receptor, has recently been identified as membrane proteoglycan (β -glycan) (Cheifetz, Andres et al. 1988). Almost all normal liver cells exhibit TGF- β 1 receptors, while the expression of the receptors is increased in activated HSCs during liver fibrosis (Friedman, Yamasaki et al. 1994).

3.2 TGF- β activation

TGF- β is secreted in a latent form with the non-covalently linked latency-associated peptide (LAP) (Barcellos-Hoff 1996). LAP is disulfide linked to another protein, latent TGF- β binding protein (LTBP), which has features in common with ECM proteins, and targets latent TGF- β to the matrix (Marek, Brodzicki et al. 2002). LAP and LTBP are required for efficient secretion of TGF- β , prevent it from binding to ubiquitous cell surface receptors, and maintain its availability in a cellular reservoir that is readily accessed by activation (Munger, Harpel et al. 1997).

Generation of active TGF- β form has been widely studied. However, the mechanism by which latent TGF- β is activated physiologically is not well understood. The activation of TGF- β is considered to involve some proteases with proteolytic activity, such as plasmin. Plasmin cleaves LAP and LTBP from the latent TGF- β and releases the active TGF- β (Rifkin, Mazzieri et al. 1999). Plasmin is derived from plasminogen by the action of urokinase plasminogen activator (uPA) or tissue

plasminogen activator (tPA) (Yee, Yan et al. 1993). Several factors can enhance PA/plasmin levels and thereby induces plasmin-mediated activation of latent TGF- β , such as retinoic acid (RA) (Okuno, Kojima et al. 2002), 9-cis-retinoic acid (Imai, Okuno et al. 1997), and Zf9 (Kojima, Hayashi et al. 2000). Moreover, De Bleser et al documented that the plasmin-mediated activation of TGF- β in HSCs was probably via the insulin-like growth factor II/mannose-6-phosphate (IGFII/M6P) receptor because IGFII/M6P receptor was upregulated in activated HSCs (de Bleser, Jannes et al. 1995).

3.3 TGF- β 1 on liver fibrosis

The critical roles of TGF- β 1 in liver fibrosis have been established by many investigations, most among which is its effects on fibrogenesis.

TGF- β 1 can cause an increase in the production of many ECM proteins, including collagen I, III, IV, fibronectin and proteoglycan (Casini, Pinzani et al. 1993). TGF- β 1 regulates gene expression through its intracellular signaling molecules such as Smad2 or Smad3. In terms of transcriptional regulation of collagen I expression, there are three protein binding sites in the promoter region of collagen I for C/EBP- β (CCAAT/enhancer binding protein- β), Sp1 and Smad3 respectively. The interaction between Sp1 and Smad3 is essential for basal and TGF- β 1 stimulated expression in HSCs (Zhang, Ou et al. 2000).

TGF- β 1 also alters the balance of matrix turnover in favor of downregulation of matrix degradation. Matrix degradation is regulated by metalloproteinases (MMPs) and their inhibitors - tissue inhibitors of metalloproteinases (TIMPs). In fibroblasts, TGF- β 1 downregulates interstitial collagenase (MMP-1), which results in a decreased degradation of collagen I. Moreover, TGF- β 1 upregulates gelatinase A (MMP-2),

which results in an increase in collagen IV (Edwards, Murphy et al. 1987; Overall, Wrana et al. 1989). Furthermore, TGF- β 1 induces the expression of TIMP-1, which inhibits proteolytic activity of all MMPs and lead to accumulation of ECM (Edwards, Murphy et al. 1987).

Due to its profound effects on ECM production and degradation, TGF- β 1 is widely considered to be the most important fibrogenic cytokine in the liver (Fig.3.). In primary HSCs culture, addition of TGF- β 1 results in an accelerated transformation to the activated phenotype (Bachem, Sell et al. 1993). Moreover, several reports have been shown that by employing soluble dominant negative TGF- β receptor or inhibitory signaling molecules could prevent liver fibrosis in experimental animals. Furthermore, TGF- β 1 knockout mice develop less liver fibrosis than normal mice (Yata, Gotwals et al. 2002).

In summary, liver fibrosis resulted from chronic hepatic damage is a character of most types of chronic liver disease. It is associated with excessive accumulation of ECM proteins and change of hepatic architecture. Recently, HSC activation has been defined as the central event of liver fibrosis. During chronic hepatic injury, HSCs undergo a phenotypic change transforming the cells from a quiescent phenotype to an activated phenotype. A number of cytokines and growth factors contribute to this activation process. Due to the profibrotic property, TGF- β 1 has been accepted as the most important factor involved in HSC activation. The advanced knowledge about the cellular and molecular mechanisms of liver fibrosis provides new therapeutic approaches to the treatment of liver fibrosis. The resolution of activated HSCs (apoptosis or convert to quiescent stage) and the anti-TGF- β strategy have become the 'hot spots' for the treatment of liver fibrosis.

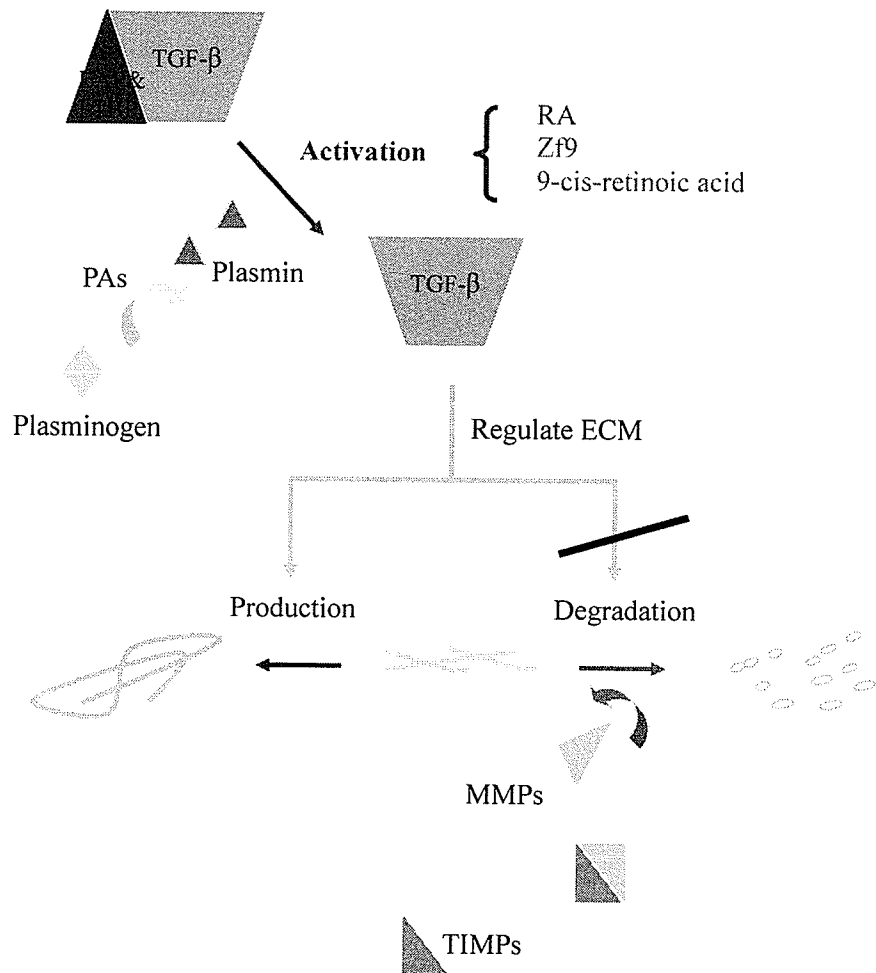


Fig.3. The effects of TGF- β on fibrogenesis. The activation of TGF- β is through the cleavage of LAP and LTBP from the latent TGF- β . This process is regulated by several cytokines including RA, Zf9 and 9-cis-retinoic acid. Active TGF- β regulates ECM turnover by increasing its production and decreasing its degradation. ECM degradation is controlled by MMPs and their inhibitor TIMPs. TGF- β induces excessive TIMPs expression to block proteinase activity of MMPs.

4. Smad Family

4.1 History

A *Drosophila* screen to pick up maternal-effect enhancers of the TGF- β -like gene, *decapentaplegic* (*dpp*), led to the identification of *Mad* (*Mother against dpp*) (Raftery, Twombly et al. 1995). Mad protein was demonstrated to function downstream of TKV, the type I receptor of dpp protein, and required for the cells receiving dpp signals (Newfeld, Chartoff et al. 1996). A *Caenorhabditis elegans* gene closely related to *Mad* was subsequently identified, which was named *Sma* (*Small*). Further mosaic analysis suggested that *Sma* gene functioned downstream of bone morphogenetic protein (BMP) type II receptor, *daf-4* (Savage, Das et al. 1996). In 1996, the first human homologue of *Mad* and *Sma* was cloned, which was later termed as Smad1 (Liu, Hata et al. 1996). Since then, eight vertebrate homologues have been found, and they are all named as Smad proteins after the first two genes identified in invertebrate (Attisano and Wrana 1998).

4.2 Subtype

Members of Smad family are functionally divided into three groups (Itoh, Itoh et al. 2000). R-Smads, receptor-activated Smads, include Smad1, Smad2, Smad3, Smad5 and Smad8, which can be phosphorylated by the type I receptors for TGF- β family members. Co-Smads, Smad4 and XSmad4 (*Xenopus*-Smad4), function as common partners, forming heteromeric complexes with the R-Smads. I-Smads, inhibitory Smads including Smad6 and Smad7, act as major negative regulators, which downregulate signal transduction of TGF- β family. R-Smads and Co-Smads are expressed in most cell types, while the expression of I-Smads is regulated by extracellular signals (Derynck and Zhang 2003).

4.3 Structure

R-Smads and Co-Smads consist of two conserved domains, Mad Homology (MH) domains I and II (Fig.4.). MH1 domains at amino terminal and MH2 domains at carboxyl terminal are separated by proline-rich linker regions of varying length. I-Smads only have MH2 domains. Their MH1 domains are replaced by divergent amino terminal, which shares similarity within the inhibitory Smad subgroup (Attisano and Tuen Lee-Hoeflich 2001).

The MH1 domain is highly related to the DNA-binding activity of certain Smads. Some R-Smads and Co-Smads can bind to a specific DNA sequence of 5'-AGAC-3', also termed the Smad-binding element (SBE) (Jonk, Itoh et al. 1998). The analysis of crystal structure of Smads demonstrates that an 11-residue β -hairpin structure in MH1 domain plays a major role in interaction with a DNA sequence. MH1 domain forms a compact globular fold which uses this β hairpin to contact the DNA in the major groove (Shi, Wang et al. 1998). Although this interaction is of relatively low specificity, it has been shown to be vital for the transcription of target genes.

The MH2 domain does not bind DNA but mediates differential association with a wide variety of proteins. Firstly, the MH2 domains of R-Smads and I-Smads have the ability to combine with the type I receptors for the members of TGF- β family (Hayashi, Abdollah et al. 1997). Activated type I receptor interact with R-Smads and phosphorylate them at their SSXS motif at the end of MH2 domain. The specificity of the interaction is determined by the L3 loop (Lo, Chen et al. 1998). Secondly, the MH2 domains of R-Smads and Co-Smads are critical for homomeric and heteromeric complex formation (Wu, Zhang et al. 1997). Finally, it has been reported that, MH1 and MH2 domains can combine with various transcription factors, controlling expression of specific genes (Moustakas, Souhelnytskyi et al. 2001).

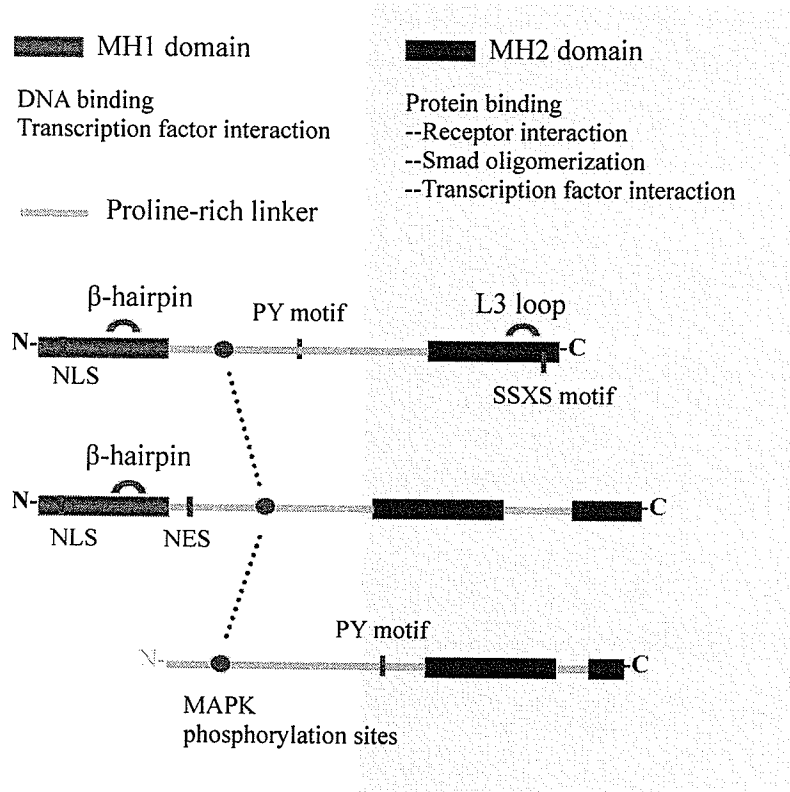


Fig.4. Schematic structures of Smads. R-Smads and Co-Smads consist of MH1 and MH2 domains, which are separated by proline-rich linker regions of varying length. I-Smads lack of MH1 domains. A number of specific sequences in these domains and linkers are functionally important for the signal transduction of TGF- β family.

The proline-rich linker region contains a number of important peptide motifs. R-Smads and I-Smads contain a conserved proline tyrosine (PY) motif that mediates the interaction with the Smad-binding proteins, Smurf-1 and Smurf-2. Smurfs are E3 ubiquitin ligases and catalyze ubiquitin-mediated degradation of certain Smads (Zhu, Kavsak et al. 1999). The linker regions also contain potential sites for phosphorylation by mitogen activated protein kinases (MAPKs). This phosphorylation can block R-Smads function (ten Dijke, Miyazono et al. 2000).

The MH1 domains and linker regions contain important sequences for nuclear translocation of R-Smads and Co-Smads. Without ligand stimulation, R-Smads localize in the cytoplasm (Fig.5) (Xiao, Watson et al. 2001). After ligand binding, R-Smads form complexes with Smad4 and then translocate into nucleus. The translocation of R-Smads is conferred by a lysine-rich nuclear localization sequence (NLS) in the MH1 domain. Importin- β binding to this NLS accumulates R-Smads in the nucleus (Derynck and Zhang 2003). Smad4 is distributed in the nucleus and cytoplasm and continuously shuttles in between. This is owing to the constitutively active NLS in its MH1 domain and a nuclear export signal (NES) in its linker region (Watanabe, Masuyama et al. 2000). This NES may be masked by R-Smads in the R-Smads/Smad4 complexes allowing the complexes to translocate into the nucleus. I-Smads are localized in the nucleus in the absence of ligand and export to cytoplasm after ligand stimulation. They do not have the indicated NLS or NES and translocate by other mechanisms (Itoh, Asao et al. 2001).

4.4 Signaling pathway

Smads are pivotal intracellular proteins that transfer signals of TGF- β family members from the cell surface to the nucleus (Fig.5.).

TGF- β family members include transforming growth factor- β s (TGF- β s), bone morphogenetic proteins (BMPs), activins/inhibins, growth & differentiation factors (GDFs) etc. They are structurally related secreted cytokines with great effects on regulating embryonic development, tissue repair and cell behavior (Massague and Chen 2000). TGF- β family members elicit signals through two types of transmembrane serine/threonine kinase receptors. Without ligand binding, type II and type I receptors exist as homodimer at cell surface. In the presence of ligand, type II receptor dimers, which contain constitutively active kinase, activate type I receptor dimers by phosphorylating the GS-domains. Activated type I receptors phosphorylate R-Smads at their SSXS motifs in a specific and transient manner. Subsequently, the phosphorylated R-Smads recruit Smad4 to form a heterotrimer or a heterodimer, which translocate into the nucleus, where they regulate target gene transcription (Derynck and Zhang 2003).

The specificity of signal transduction is determined by type I receptors (Massague 1998). In mammals, five type I receptors and seven type II receptors have been identified, which also name as activin receptor-like kinases (ALKs), (Fig.6). The BMP type II receptor (BMP-RII) can combine with three type I receptors, ALK-2 (ActRI), ALK-3 (BMP-RIA), and ALK-6 (BMP-RIB), mediating BMP signaling through Smad1, Smad5 and Smad8. ALK-2, initially implicated as a type I receptor for TGF- β s, is now considered as the most important receptor for BMP signaling. On the other hand, TGF- β type II receptor (T β RII) interact with type I receptor ALK-5 (T β RI), which activates Smad2 and Smad3. Recently, ALK-1, which transfers signals through Smad1 and Smad5, has been identified as specific TGF- β type I receptor in endothelial cells (Itoh, Itoh et al. 2000).

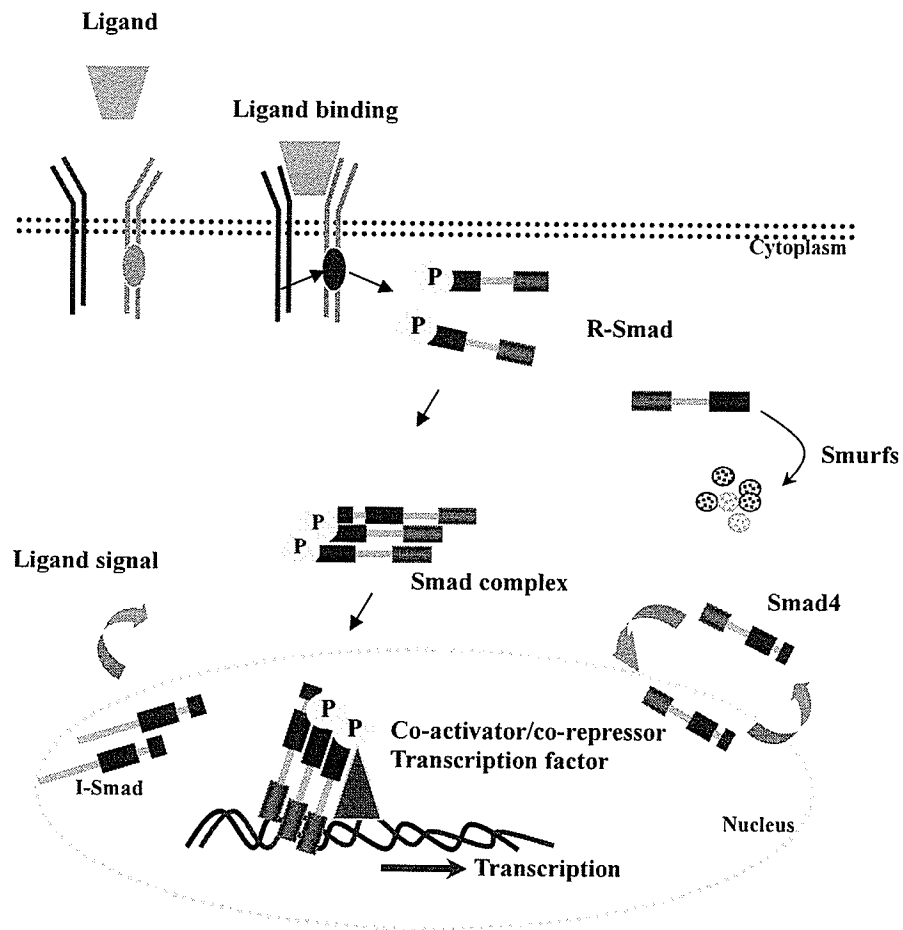


Fig.5. Signaling pathway of TGF- β family. With ligand binding, type II receptor dimers activate type I receptor dimers, which subsequently phosphorylate R-Smads. Activated R-Smads form a heterotrimer with Smad4. This Smad complex translocates into the nucleus, which regulate responsive gene expression by association with co-activator/co-repressor or other transcription factors. Without ligand stimulation, R-Smads are localized in the cytoplasm while I-Smads are localized in the nucleus. Smad4 is distributed in the nucleus and cytoplasm and continuously shuttles in between. Smurfs interact with certain R-Smads and target them for degradation.

Type II receptors	Type I receptors	R-Smads
BMPRII	ALK-2 (ActRI) ALK-3 (BMP-RIA) ALK-6 (BMP-RIB)	Smad1, 5, 8
T β RII	ALK-5 (T β RI) ALK-1 ALK-2	Smad2, 3 Smad1, 5

Fig.6. Receptors and intracellular mediators of TGF- β family. TGF- β and BMP signaling pathways are listed. BMPRII can interact with three type I receptors (ALK-2, ALK3 and ALK-6) transferring signals through Smad1/5/8. T β RII activate ALK-5, ALK-1 or ALK-2 to phosphorylate Smad2/3 and Smad1/5.

In the nucleus, Smads modulate gene expression through association with transcriptional co-activators/co-repressors. The structurally related p300 and CBP (CREB-binding protein) are the first two co-activators identified in the study of Smads transcriptional activity (Feng, Zhang et al. 1998). p300/CBP are shown to interact with phosphorylated R-Smads through binding to their MH2 domains and promote transcription by loosening the chromatin structure. Smads have also been found to interact with transcription factors. An example is the regulation of *Xenopus Vent-2* gene (Hata, Seoane et al. 2000). OAZ, which belongs to a zinc finger transcription factor family, is found to cooperate with Smad1 and Smad4 in BMP2-induced activation of *Xenopus Vent-2* promoter.

The signal transduction of TGF- β family is controlled by the negative effects of I-Smads and Smurfs. I-Smads downregulate the signal transduction through competing with R-Smads for receptors or co-Smads interaction (Hayashi, Abdollah et al. 1997). I-Smads can interact efficiently with the activated type I receptors or co-Smads, thereby preventing the activation of R-Smads or the formation of R-Smads/Smad4 complexes. Other than Smad7 acting as a general inhibitor of TGF- β family signaling, Smad6 appears preferentially inhibit BMP signaling pathway (Ishisaki, Yamato et al. 1999). In Addition, induction of I-Smads expression by TGF- β family members represents an auto-inhibitory feedback mechanism for the signal transduction (Massague 2000). Smurfs (Fig.5), including Smurf-1 and Smurf-2, also antagonize TGF- β family signaling (Zhu, Kavsak et al. 1999). Smurfs (Smad ubiquitination regulatory factors) are members of the HECT (homologous to the E6-AP carboxy terminus) subclass of E3 ubiquitin ligases. Smurfs interact with certain R-Smads through binding to their PY motifs in the linker regions and target them for degradation. Thus, Smurf-mediated degradation controls R-Smads level and the sensitivity of cells to incoming signals. It has been reported that I-Smads also contain the binding sites for Smurfs (Murakami, Watabe et al. 2003).

5. Smad1 Gene

5.1 Comparison of Smad1 gene

Smad1 gene in human (Riggins, Thiagalingam et al. 1996), mouse (Zhao and Hogan 1997) and rat (Chen, Lebrun et al. 1996) have seven exons interrupted by six introns. The first exons are noncoding exons (Fig.7).

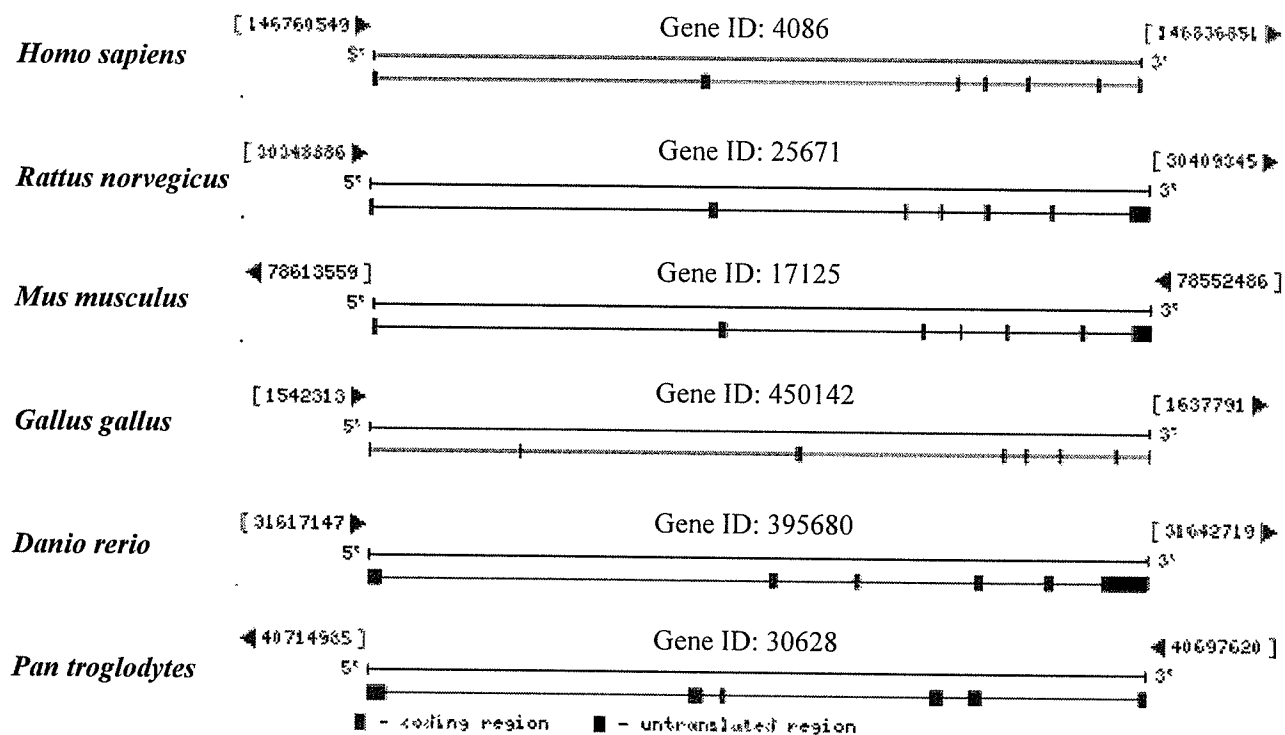


Fig.7. Smad1 gene in different species. Human, mouse and rat Smad1 have 7 exons, interrupted by 6 introns. The Gene information was obtained from <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene&cmd=search&term=smad1>.

The 5'-noncoding region of Smad1 is conserved in different species and may play an important role in modulating the gene expression at translational level. Human and mouse Smad1 proteins are identical at the length, both having 465 amino acids (Liu, Hata et al. 1996; Strausberg, Feingold et al. 2002). Rat Smad1 is little different in sequence, carrying 468 amino acids (Yue, Hartsough et al. 1999).

5.2 Rat Smad1 gene

Two types of rat Smad1 cDNA sequences were found in Genbank database: a sequence of 1425bp (Genbank accession no. AF067727) and a sequence of 3010bp (Genbank accession no. NM_013130). The alignment analysis (Fig.8) revealed that both sequences contains entire open reading frame (ORF) coding a protein of 468 amino acids. The longer one includes extra sequences at both ends of the cDNA representing the 5'-UTR (untranslated region) and 3'-UTR. Since the additional sequences are noncoding, the two transcripts are translated into the same protein.

According to different cDNA sequences, the genomic DNA organizations of rat Smad1 are different (from Ensembl database). For cDNA of 1425bp, Smad1 gene spans 30.4 kb of the rat Chromosome 19 (Fig.9) and contains six exons interrupted by five introns. Exon 1 and exon 6 include a fragment of untranslated region respectively. On the other hand, for cDNA of 3010bp, the rat Smad1 gene spans 60 kb of the rat genome and contains seven exons separated by six introns. The first exon is noncoding, while other six exons contain a short 5'-UTR, the entire 1407bp ORF, and a 3'-UTR (Fig.10). Since more and more evidence supports for the longer transcript, cDNA of 3010bp and its genomic organization of seven exons are accepted as rat Smad1 transcript and genomic DNA structure at present.

Rat Smad1 transcript I

Genbank No. AF067727 (1425bp)

Rat Smad1 transcript II

Genbank No. NM 013130 (3010bp)

Rat Smad1 protein

Genbank No. AAC19116 (468aa)

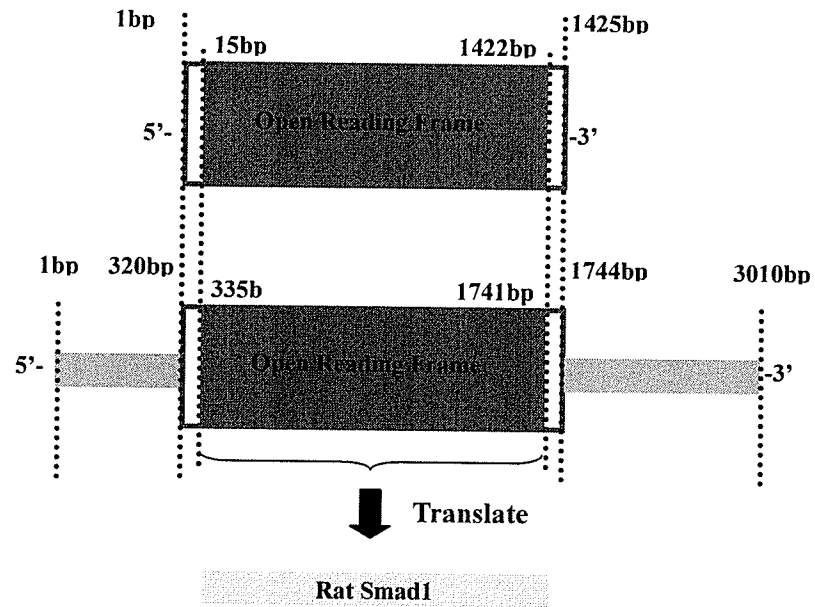


Fig.8. Alignment analysis of two rat Smad1 transcripts. Alignment was performed by online tool from

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi>

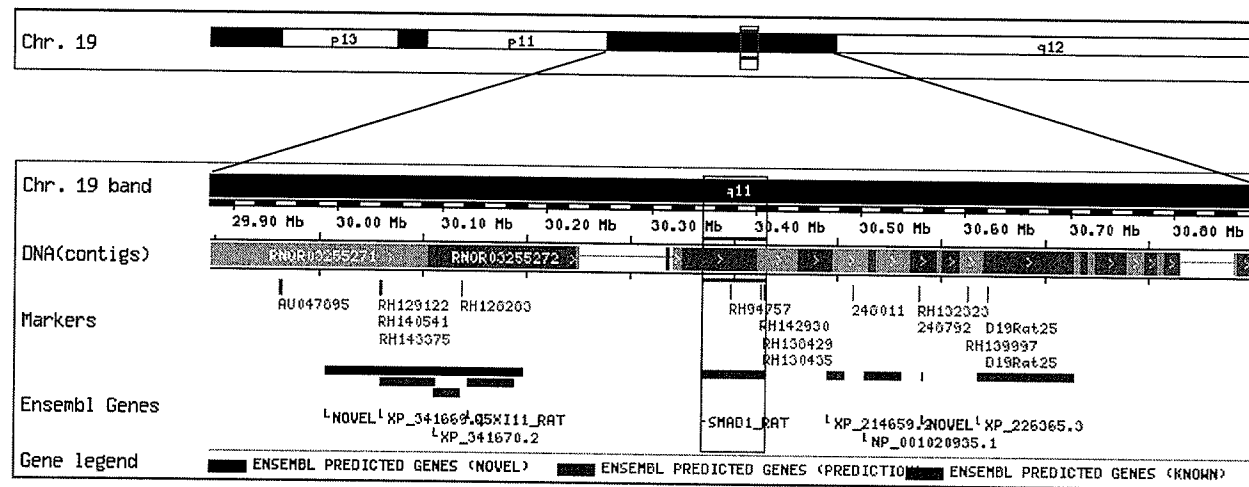


Fig.9. An overview of rat Chromosome 19. Figures from Ensembl website:

http://www.ensembl.org/Rattus_norvegicus/contigview?l=19:30348886-30409345

Transcript of 1425bp in genomic location

Transcript of 3010bp in genomic location

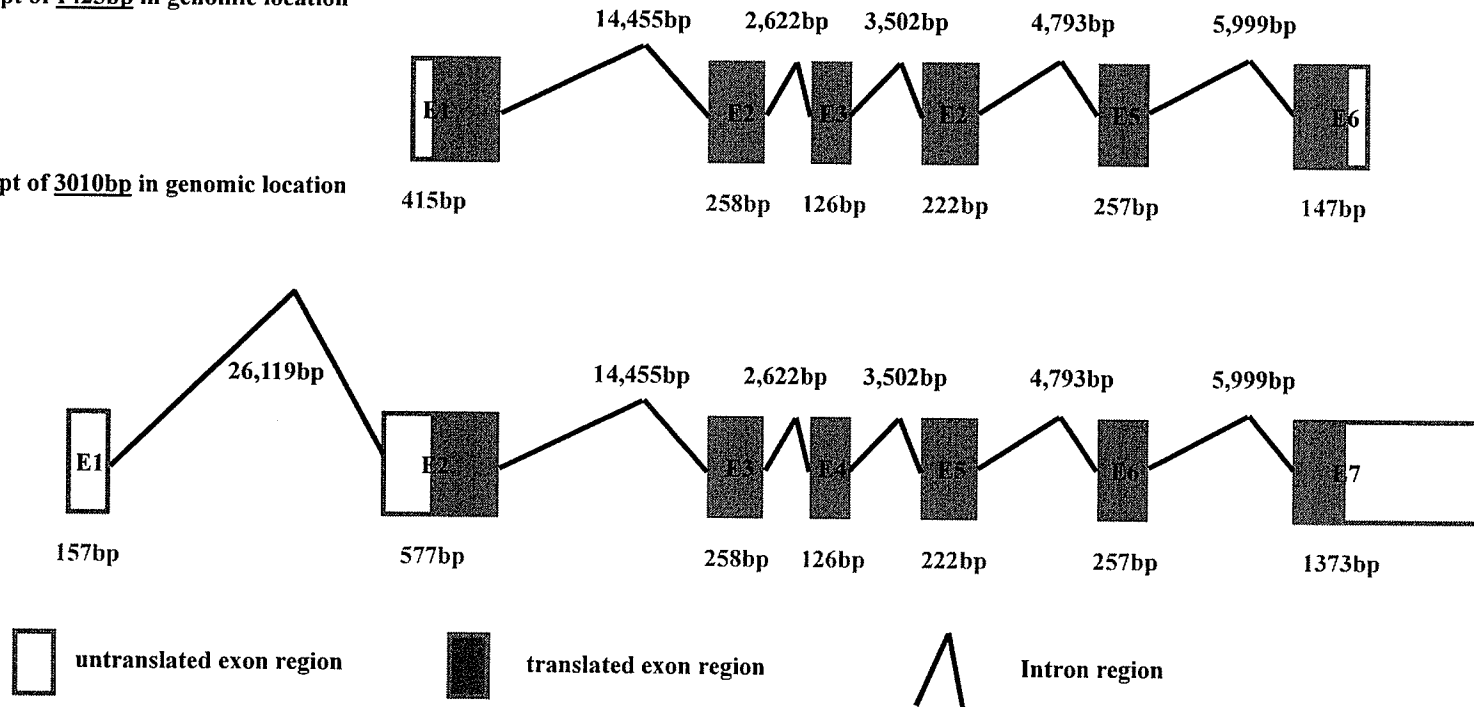


Fig.10. Rat Smad1 gene information from Ensembl website:

http://www.ensembl.org/Rattus_norvegicus/exonview?db=core;transcript=ENSRNOT00000025079.

6. Gene Expression Regulation

6.1 Regulation of gene expression

Gene expression refers to a process that a DNA sequence is firstly transcribed into RNA and then produced into a functional protein. RNA synthesis from a gene is transcription and the protein synthesis from the RNA is translation (Cooper 1981).

There are many regulatory steps in the pathway from DNA to protein, which are critical to regulation of gene expression (Shames, Selzman et al. 1998). At transcriptional level, the regulation can be at the activity of transcription, such as time and level of a given gene expression. At post-transcriptional level, the expression of a gene can be controlled by either splicing of pre-RNA or RNA degradation. RNA degradation could determine the fate of certain mRNA molecules in the cytoplasm. At translational level, protein production can be regulated by either translational rate or degradation of the protein. Although there are several regulatory steps in gene expression, the transcription of a gene is considered to be the most important step (Brown 1982).

6.2 Control of transcriptional activity

Transcription is the process that a DNA sequence is enzymatically copied by a RNA polymerase to produce a complementary RNA, thus transfers the genetic information from DNA into RNA. Transcription initiates with RNA polymerase and its associated accessory proteins recognizing and specifically binding to the DNA sequence. In eukaryotes, three RNA polymerases have been identified. RNA polymerase II transcribes protein-encoding genes into mRNA (Shenk 1981).

Control of transcriptional activity involves several elements in the 5'-flanking region of genes (Fig.11) (Etienne, Brault et al. 1990). These elements include

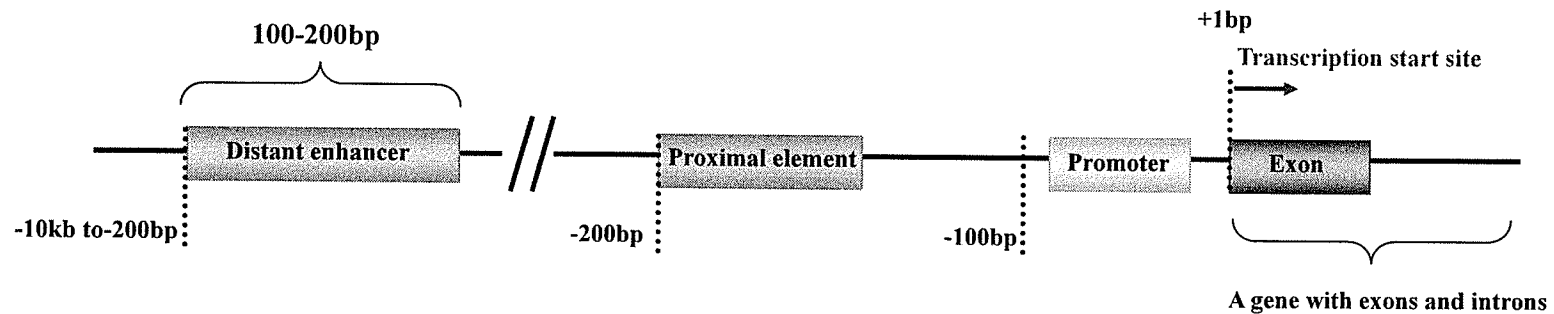
5'-flanking region of a gene

Fig.11. 5'-flanking region of a gene. Transcription of a gene is controlled by several components in 5'-flanking region, including a transcription start site, a promoter, proximal elements and distant enhancer sites. A promoter is usually located within 100bp upstream from transcription start site. Promoter-proximal elements occur within ≈ 200 bp of the start site, and enhancers, often 100-200bp in length, may be located distantly from a gene.

transcriptional start site (TSS), promoter, promoter-proximal elements and enhancers. TSS is where transcription of the gene into RNA starts. Promoter is the binding site for transcription-initiation complex. Transcription-initiation complex consists of a RNA polymerase and several general transcription factors. RNA polymerase is the enzyme triggers transcription reaction and general transcription factors are a series of proteins that help RNA polymerase locate to the proper site in a DNA template. These factors are consistent in all cell types. Therefore, they are referred as 'general' transcription factors. Promoter-proximal elements and enhancers are generally same, but we use the term 'promoter-proximal' for the elements lying within 100-200bp upstream of the TSS, and 'enhancer' for the elements that regulate gene expression distantly. Analysis of different eukaryotic enhancers has shown that enhancers can locate before, after, or inside the gene. Once transcription factors bound to the enhancer, the DNA sequence between promoter and enhancer loops out to allow the factors interact with initiation complex and then target it assemble at the promoter. The transcription factors bound to promoter-proximal elements and enhancers are cell specific and in most cases they stimulate gene transcription. Among these elements, promoter is the most important. The sequence of promoter determines the binding of transcription-initiation complex to the DNA template and the basal transcriptional activity of the gene. Moreover, assembly of the initiation complex on the promoter can usually occur in the absence of the upstream elements. Furthermore, the 5'-flanking region of gene is also generally called promoter region.

6.3 Types of promoters

The term of promoter was first used by geneticists in 1964 to describe the function of a locus immediately upstream of the three genes in the lactose *operon* (Travers 1974). When the locus was mutated, the genes in the operon were not expressed. Therefore, the locus appeared to promote the gene expression. Now, promoter generally refers to the binding site of transcription-initiation complex.

In eukaryotic cells, the 'classical' promoters for RNA polymerase II contain a TATA box. TATA sequence was first discovered in *Escherichia coli* promoters. *E. coli* promoters consist of two segments of (5'-TTGAACA-3') and (5'-TATAAT-3') located at 35bp and 10bp upstream of the TSS respectively. Both of these segments are important for gene transcription. In eukaryotes, TATA box has been found in the promoter region of many rapidly transcribed genes. These genes are very actively transcribed either at particular times of the cell cycle or in specific differentiated cell types. The TATA box of a consensus sequence (5'-TATAWAW-3'), where W represents A or T, is usually found at 25–35bp upstream of the TSS. Mutagenesis studies have shown that a single-base change in this nucleotide sequence drastically decreases the *in vitro* gene transcription (Benoist and Chambon 1981).

In addition to the TATA box, the promoters of many other genes transcribed by RNA polymerase II contain another important sequence element named initiator (Inr). Unlike the highly conserved TATA box sequence, only a degenerated consensus sequence for Inr has been defined, which is 5'-YYANT/AYYY-3'. In this sequence, A is usually the start site of TSS, Y could be either C or T, and N could be any of the

four bases. Mutagenesis studies have revealed that the strength of the Inr promoters is determined by the nucleotide sequence immediately surrounding the start site. Moreover, some promoters for RNA polymerase II contain only an Inr element without TATA box. In such TATA-less promoters, the Inr element is critical in positioning RNA polymerase II (Weis and Reinberg 1992).

Other types of promoter have been identified. In stead of a TATA box or an Inr element, a CG rich sequence of 20-50bp in length is found in some genes (Tamura and Mikoshiba 1991). Unlike the gene with TATA or Inr promoter, a gene containing a CG rich promoter begins its transcription at multiple start sites, which extend from 20 to 200bp in length, therefore, resulting in the production of mRNAs with multiple alternative 5'-ends. This CG rich promoter is usually located within 100bp upstream of the transcription start site. Genes containing CG promoter are generally described at low rate and encode the enzymes of intermediary metabolism such as the house-keeping genes. The corresponding binding protein of CG rich promoter is Sp1 (specificity protein 1). Sp1 is a transcription factor that recognizes the sequence 5'-GGGCGG-3' in the promoter region (Fig.12).

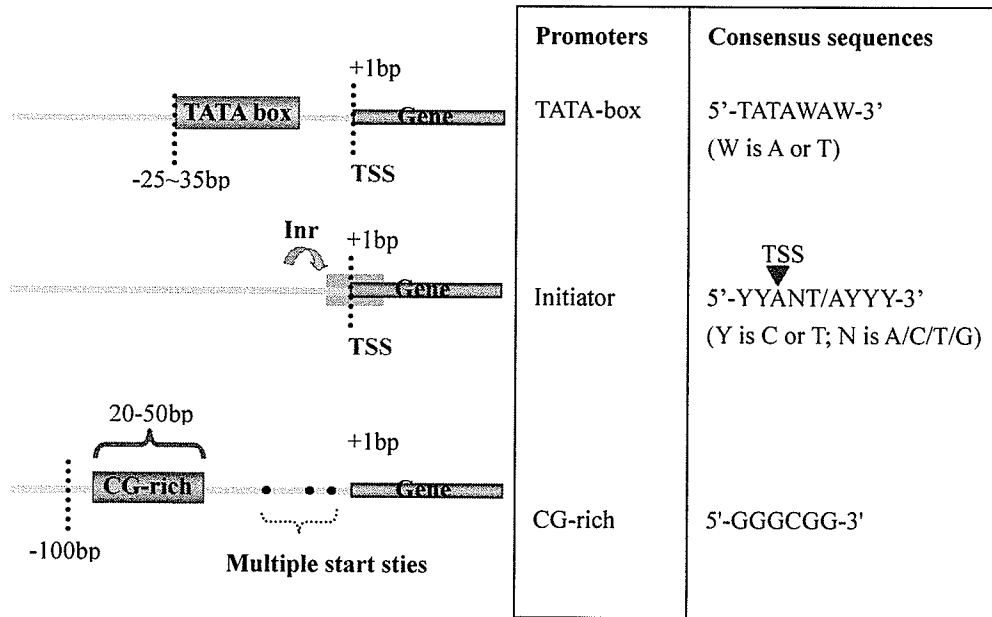


Fig.12. Comparison of three types of eukaryotic RNA polymerase II promoters. Three types of promoter for RNA polymerase II are identified, which include TATA box promoter, Inr promoter and CG-rich promoter.

7. Previous Studies

7.1 *Smads and liver fibrosis*

Due to the important role of TGF- β on fibrogenesis, the downstream mediators Smad2, Smad3, Smad4 and Smad7 have been widely studied by many researchers.

Smad2 and Smad3 are the signaling molecules of TGF- β s. After activation by TGF- β type I receptors, they translocate into the nucleus and regulate the corresponding gene expression. It has been reported that the role of Smad2 and Smad3 in liver fibrosis is different (Liu, Gaca et al. 2003). Smad2 activation and

nuclear translocation occurs primarily in quiescent hepatic stellate cells, whereas Smad3 activation occurs primarily in activated HSCs. The role of Smad3 in HSC activation is demonstrated by Smad3 knockout mice treated with carbon tetrachloride (CCl₄) (Schnabl, Kweon et al. 2001). Smad3 is required for collagen I expression and TGF- β induced HSC proliferation. Inhibition of Smad3 decreases the fibrotic degree in the animal model (Zanninelli, Vetuschi et al. 2006).

Smad4, which is the common partner of R-Smads, is required by transduction. Activated Smad2 or Smad3 forms a heteromeric complex with Smad4 and transfers TGF- β signals in to the nucleus. Inhibition of Smad4 blocks TGF- β signaling. Antisense Smad4 gene was generated to explore the effect of Smad4 on HSC activation (Xu, Leng et al. 2004). Loss function analysis revealed that antisense Smad4 can inhibit HSC activation and proliferation. Additionally, the production of extracellular matrix was obstructed as well. Recently, Smad4 has been considered as a studying target for the development of antifibrotic gene therapy.

Smad7 is an intracellular antagonist of TGF- β signaling. Overexpression of Smad7 inhibits the collagens production and HSC proliferation and causes HSCs remain in quiescent stage, which ameliorates liver fibrosis (Dooley, Hamzavi et al. 2003).

7.2 *Smad1 and liver fibrosis*

Smad1, which transmits signals for both BMPs and TGF- β s, has not been notable in the research field of liver fibrosis until recently (Shen, Huang et al. 2003).

Previous studies in our laboratory demonstrated that the expression of Smad1 increased dramatically during the activation of HSCs, which were isolated from rat liver and activated by prolonged culture on uncoated plastic dishes. The activation process was monitored by the expression of α -SMA (α -Smooth muscle actin), a marker of activated HSCs. Total proteins on days 1, 3, 6, 9 and 12 were collected and the expression of Smad1 was detected by specific antibody. We found that the obvious increase in Smad1 was observed at day 6 of primary cultured HSCs. Similar increasing pattern was observed on the mRNA level by employing both classic RT-PCR and quantitative real-time PCR.

We also found that the enhanced expression of Smad1 during HSC activation was inhibited by addition of TGF- β 1. Rat HSCs were cultured for six days, and then TGF- β 1 was added and incubated for different hours. The abundance of Smad1 mRNA was reduced in a consistent manner. Further experiment was carried out to investigate the expression of Smad1 in response of different amount TGF- β 1 administration. HSCs after six days of primary culture and incubated with different amount of TGF- β 1 for 12 hours. The abundance of Smad1 mRNA and protein was decreased with an increased TGF- β 1 concentration.

These findings revealed the possible role of Smad1 in liver fibrosis. Unraveling the control mechanism of Smad1 expression will contribute to the understanding of the complex network of gene regulation in HSC activation. Therefore, it will eventually be possible to identify the factors that are important during hepatic fibrogenesis.

Hypotheses:

- Since there is an increase in Smad1 expression during the activation of hepatic stellate cells, we hypothesize that there are different promoter activities of Smad1 in Quiescent and Activated HSCs.

Objectives:

- To clone of the promoter of rat Smad1 gene
- To investigate the transcriptional activity of rat Smad1 promoter in quiescent and activated hepatic stellate cells.
- To analysis the binding transcriptional factors

II. MATERIALS AND METHODS

1. Materials

Unless specified elsewhere, all the chemicals were purchased from Sigma-Aldrich. The cell culture mediums and serum were purchased from VWR.

Table1. Materials

Product	Company	Experiment
100mM dNTP set, PCR Grade	Invitrogen	PCR reactions purification
BD Matrigel™	BD Biosciences	Cell culture
DH5™ Competent Cells	Invitrogen	Chemical transformation
DISPASE	BD labware	Cell culture
Dual-luciferase reporter assay system	Promega	Luciferase Assay
Dulbecco's Modified Eagle Medium	GIBCO/BRL	Cell culture
Gamma 32p-ATP	GE health care	DNA labeling
GFX PCR DNA and Gel Band Purification Kit	GE health care	Gel purification
GLprimer1, GLprimer2	Promega	DNA sequencing
Kodak Biomax XAR film	GE health care	Autoradiography
Lipofectamine2000	Invitrogen	Transient Transfection
MicroSpin G25 column	GE health care	Labeled DNA purification
Non-essential amino acids	GIBCO/BRL	Cell culture
Nuclei EZ prep kit	Sigma-Aldrich	Nuclear protein extraction
Nucleic Acid Marker	Invitrogen	DNA electrophoresis
pCR-Topo4	Invitrogen	DNA Subcloning
Penicillin, Ampicillin	GIBCO/BRL	Cell culture and bacteria culture
pGL2-Basic vector,	Promega	DNA Subcloning
Polyd(I-C)	GE health care	EMSA
pRL-TK vector	Promega	DNA Subcloning
Protein assay kit	Bio-Rad	Protein conc. measurement
Proteinase inhibitor	Invitrogen	Protein storage
Proteinase K	Invitrogen	Genomic DNA isolation
QIAGEN Plasmid Mini Kit	Qiagen	Plasmid DNA isolation
QIAquick PCR Purification	Qiagen	PCR reactions purification
Restriction enzymes	Promega	Restriction enzyme digestion
RNase, DNase-free	Roche	Plasmid DNA isolation
T4 DNA ligase	Promega	DNA ligation
T4 Polynucleotide Kinase	Promega	DNA labeling
Taq DNA Polymerase	Qiagen	Polymerase Chain Reaction

2. Methods

2.1 Genomic DNA Isolation

Kidney tissue from adult Sprague-Dawley rats was kindly provided by Ms. Jianghong Fan (University of Manitoba, Canada). Approximately 0.3g of the tissue was homogenized (Brinkman Homogenizer, Brinkman Instrument, Westbury, NY) and digested in 1ml Proteinase K solution (100mM Tris-Cl, pH 8.0, 100mM NaCl, 5mM EDTA, 1% SDS and 100ug/ml Proteinase K) overnight (O/N) at 37°C (Precision automatic CO₂ incubator, Precision Scientific Inc, Illinois).

Equal amount of buffered phenol and chloroform mixture (v/v, 1:1) was added to the solution and mixed by inverting the tube for several times. After centrifugation for 5 minutes at 12,000rpm (MicroMax centrifuge, Thermo IEC, Waltham, MA), aqueous layer was transferred to a new tube. Repeated this step 2~3 times, until aqueous layer was clear and no protein interface was visible. Repeated this step one more time with chloroform and transferred the supernatant to a new tube.

Genomic DNA was precipitated with proper amount of 3M NaOAC and 100% ethanol at -70°C for 2~3 hours and centrifuged at 15,000rpm for 15 minutes. The amount of NaOAC and ethanol was calculated by the following formula.

$$V_{\text{NaOAC}} = 10\% * V_{\text{aqueous layer}}$$

$$V_{\text{100\% ethanol}} = 2 * (V_{\text{aqueous layer}} + V_{\text{NaOAC}})$$

Discarded the supernatant and washed DNA pellet with 70% ethanol twice. DNA pellet was air dried for 15 minutes and resuspended in proper amount of TE buffer (10 mM Tris-Cl, pH 8.0 and 1 mM EDTA). Concentration of genomic DNA

was measured by a spectrophotometer (Bio-Rad SmartSpec 3000 Spectrophotometer, Life Science Research Division, Mississauga, ON) and the quality was checked by restriction enzyme digestion.

2.2 Restriction Enzyme Digestion

DNA was digested with various restriction endonucleases (Promega, Madison WI). Restriction enzyme digestion was performed in a volume of 20 μ l, which consisted of 0.2-1.5g of substrate DNA, 0.5-1 μ l of 10u/ μ l restriction enzyme, 2 μ l of 10X restriction enzyme (RE) buffer, 0.2 μ l of 10 μ g/ μ l acetylated BSA, and sterile distilled water. Digestion reaction was incubated at 37°C (Model 202 Water Bath, Fisher Scientific) for 1~4 hours and checked by Agarose Gel Electrophoresis.

2.3 Agarose Gel Electrophoresis

DNA fragments were analyzed by Agarose Gel Electrophoresis. Agarose powder was mixed with 1X electrophoresis buffer TBE (1L of 10X TBE buffer contains 108g of Tris base, 55g of boric acid and 9.3g of Na₄EDTA) to desired concentration (0.8%-2% according to the size of DNA fragments), and then heated in a microwave oven (WaveplusII Microwave Oven, GoldStar) until completely melted. For later visualization, ethidium bromide (final concentration of 0.5 μ g /ml) was added to the gel solution. Gel solution was poured into a casting tray (Midi-horizontal electrophoresis unit, Bio-Rad, USA) containing a sample comb, and allowed to solidify at room temperature.

After the gel was solidified, the comb was taken out and the gel was covered with 1X TBE buffer. DNA samples mixed with loading buffer (10X loading buffer: 0.2ml of 12.5% bromophenol blue, 0.2ml of 12.5% xylene cyanol, 25g of Ficoll® 400 and sterile water per 100ml) were subsequently pipetted into sample wells and DNA marker (Invitrogen Canada Inc, Burlington, ON) was loaded in adjacent wells. The electrophoresis was performed at 100v until DNA fragments were separated. Gel was checked on an ultraviolet transilluminator (312nm Transilluminator, Fisher Scientific, PA) and pictures were taken by an instant camera (Direct Screen instant camera, Polaroid).

2.4 PCR Cloning

Polymerase Chain Reaction (PCR) primers were designed by Oligo5 program according to the DNA sequence and synthesized by Invitrogen (Invitrogen Canada Inc, Burlington, ON). PCR was carried out in a total volume of 50µl, containing 50-100ng of template DNA, 10µl of 5X Q-Solution, 5µl of 10X QIAGEN PCR buffer, 2µl of 25mM MgCl₂ solution, 2.5 units of Taq DNA polymerase (Qiagen, Mississauga, ON), 1µl of 10mM dNTP mix (Invitrogen), 1µl of 10mM specific primers, and sterile distilled water. PCR amplification was performed using an Eppendoff MasterCycler (Eppendoff, Westbury, NY). The cycling conditions were as follows: initial denaturation at 95°C for 5 minutes, followed by 29 to 32 cycles of denaturation at 95°C for 1 minute, annealing for 45 seconds at different annealing temperatures listed in table4 and 7, extension at 72°C for 90 seconds, and then followed by 10 minutes of

final extension at 72°C. PCR products were analyzed by electrophoresis.

2.5 Reconstruction of plasmid

For cloning of rat Smad1 the 5'-flanking region and intron1 region, PCR primers were designed according to rat genomic DNA sequence from Ensembl. PCR was carried out with genomic DNA isolated from rat kidney tissue and a series of Smad1 primers with restriction enzyme sites at their ends for later subcloning. After reaction, PCR products were precipitated by 3M NaOAC and 100% ethanol at -70°C O/N and redissolved in 20µl of sterile distilled water to remove salts.

pGL2-Basic vector (Promega) containing firefly luciferase gene without promoter as a reporter plasmid was purchased from Promega and used for reconstruction of plasmid. PCR products (the 5'-flanking region and intron1 region of rat Smad1) and pGL2-Basic vector were incubated with appropriate restriction enzymes O/N for complete digestion. Target DNA fragments were purified from an Agarose gel. Digested PCR fragments and linearized pGL2-Basic plasmid with cohesive ends were joined together by DNA ligation to generate rat Smad1.pGL2 constructs.

a) DNA Gel Purification

DNA fragment was purified using GFX PCR DNA and Gel Band Purification Kit (GE health care, Mississauga, ON) from an Agarose gel. Gel slices containing target DNA were incubated in 300µl of Capture buffer at 65°C for 10 minutes until melted completely. Gel mixture was applied on a GFX column, incubated at room

temperature for 2 minutes, and then centrifuged at 15,000rpm for 30 seconds. Flow-through was discarded and column with DNA binding was washed by 500µl ethanolic wash buffer. DNA was then eluted with 30µl of 10mM Tris-Cl, pH 8.0 by centrifuging at 15,000rpm for 1 minute.

b) DNA Ligation

Two DNA fragments were ligated by T4 DNA ligase (Promega). The amount of the fragments for sticky end ligation was calculated by this formula:

$$\text{Amount}_{\text{insert}} (\text{ng}) = \frac{\text{Length}_{\text{insert}} (\text{in bp})}{\text{Length}_{\text{plasmid}} (\text{in bp})} * 2-3 \times \text{amount of plasmid (ng)}$$

Ligation reaction contained 200ng of linearized plasmid digested to produce compatible ends with the insert fragment, proper amount of insert fragment, 2µl of 10X T4 ligase buffer, 1-6 units of T4 ligase and sterile distilled water in a total volume of 20µl. The reaction was performed at room temperature O/N and then used in bacterial transformation.

2.6 Bacterial Transformation

Competent cells used in bacterial transformation were prepared from DH5α cells (Invitrogen). Proper amount of competent cells solution (usually 100µl) and 20µl of ligation reaction were mixed and incubated on ice for 30 minutes. This mixture was subjected to heat shock by immersion in a 37°C water bath for 45 seconds, and then put on ice for another 2 minutes. One milliliter of pre-warmed LB medium was

added to the solution and incubated in a 37°C shaking incubator (C₂₄ incubator shaker, New Brunswick Scientific, NJ) for 1 hour. The mixture was spun down by centrifuging at 1,000rpm for 10 minutes and resuspended gently in 100µl of LB medium. This transformation was spread on a pre-warmed LB-agar plate containing appropriate antibiotic (100µg/ml Ampicillin) and the plate was incubated at 37°C O/N for bacteria growth (Precision * incubator, Precision Scientific Inc., Chicago, IL).

2.7 Plasmid DNA Isolation

A single bacterial colony was transferred into 2-5ml of LB medium containing 100µg/ml of Ampicillin and incubated O/N at 37°C with vigorous shaking at 300rpm.

Bacteria were pelleted at 12,000rpm for 30 seconds. Pellet was resuspended in 100µl of ice-cold lysis buffer (50mM glucose, 25mM Tris-Cl, pH 8.0, and 10mM EDTA pH 8.0) by vigorous vortexing and then incubated at room temperature for 5 minutes. Alkaline solution (0.2N NaOH and 1% SDS) was freshly prepared, 200µl of which was added into the lysate and mixed by inverting the tube rapidly for 5 times. The mixture was incubated at room temperature for another 5 minutes and then 150µl of ice cold solution (60ml of 5M potassium acetate, 11.5ml of glacial acetic acid and 28.5ml of sterile distilled water) was added. This mixture was vortexed gently in inverted position for 10 seconds and cultured on ice for 5 minutes.

After centrifugation at 12,000rpm for 5 minutes, supernatant was transferred to a new tube. Equal amount of phenol/chloroform was added to extract DNA. After centrifuged at 12,000rpm for 2 minutes, aqueous layer was transferred to a new tube

and DNA was precipitated with 2 volumes of 100% ethanol. DNA pellet was washed with 70% ethanol twice, air dried for 15 minutes and redissolved in 20µl of TE buffer containing 0.2µl of 100µg/ml DNase-free RNase (Roche Mississauga, ON). Positive clones were selected by restriction enzyme digestion and results were confirmed by DNA sequencing (University Core DNA & Protein Services, University of Calgary, Canada).

2.8 Cell Culture and morphology

Human transformed liver cell line (Chang liver cells) and rat hepatoma cell line (H-4-II-E cells) were purchased from American Type Culture Collection (ATCC). Two rat hepatic stellate cell lines, CFSC-8B (cirrhotic fat-storing cell-8B) and HSC-T6 (Hepatic stellate cell-T6), were kindly provided by Dr. Marcos Rojkind (The George Washington University, DC) and Dr. Scott. L. Friedman (Mont Sinar School of Medicine, NY) respectively. Cells were grown in a 37°C humidified chamber (Water-Jacketed Incubator, Fisher Scientific) containing 5% CO₂. Plating density for each cell line was 1×10^7 per 10mm-dish. Culture medium (Table2) was routinely changed every 3-4 days.

CFSC-8B cells were also seeded on MatrigelTM dishes (MatrigelTM, Becton Dickinson Labware, MA), using the 'thin gel' method (50µl cm⁻² of growth surface) according to the manufacturer's instructions and harvested with Dispase (Becton Dickinson Labware).

Cultured cells were imaged and photographed on an Olympus inverted-phase

Table2. Cell culture conditions

Name	Type	Culture Medium	Source
CFSC-8B	Rat hepatic stellate cells	HyQ MEM/EBSS supplemented with non-essential amino acids, 10% FBS, penicillin (100U/ml) and streptomycin (100µg /ml).	Dr. Marcos Rojkind The George Washington University, Washington DC
Chang liver	Human transformed liver cells	DMEM supplemented with 10% FBS, penicillin (100U/ml) and streptomycin (100µg /ml).	ATCC
H-4-II-E	Rat hepatoma cells	HyQ DME/F-12 supplemented with 10% FBS, penicillin (100U/ml) and streptomycin (100µg /ml).	ATCC
HSC-T6	Rat hepatic stellate cells	DMEM supplemented with 10% FBS, penicillin (100U/ml) and streptomycin (100µg /ml).	Dr. Scott L. Friedman Mont Sinar School of Medicine, NY

microscope (CK-40) using a mounted Olympus 35-mm camera (Carsen Group Inc. Markham, ON) and Kodak 35mm print films (Eastern Kodak Co., Rochester, NY).

2.9 Transient Transfection

Chang liver cells and CFSC-8B cells were transfected with Lipofectamine 2000 (Invitrogen). Each 35mm-dish was seeded with $1-5 \times 10^7$ cells and the cells were grown for 1-2 days to 90% confluence. Cells were starved by culturing in 1.5ml of transfection medium (medium without antibiotics and serum) for 2-4 hours before transfection.

For each well, 4 μ g of DNA was diluted in 250 μ l of transfection medium and 10 μ l of Lipofectamine 2000 reagent was diluted in 250 μ l of transfection medium and incubated at room temperature for 5 minutes. Diluted DNA was combined with diluted Lipofectamine 2000 reagent and incubated together for 20 minutes at room temperature to form DNA/Lipofectamine 2000 complexes.

DNA/Lipofectamine 2000 complexes were added to each well containing cells and 1.5ml of transfection medium. Transfected cells were incubated in a 5% CO₂ incubator O/N, and then propagated in regular culture medium for another 24-48 hours until they were ready to assay for Dual Luciferase Assay.

2.10 Protein Isolation

Cells washed with PBS buffer and harvested by trypsinization (Trypsin, VWR Scientific, Baltimore, MD) at 37°C for 2-5 minutes. After trypsinization, cells were collected by centrifuging at 15,000rpm for 1 minute. Cell pellet was resuspended in 60-100 μ l of lysis buffer (0.1M Tris-Cl pH=7.8 and 0.1% TritonX-100) and incubated on ice for 15 minutes. The lysate was spun down at 4°C for 15 minutes. Supernatant

was collected and protein concentration was determined using a protein assay kit (Life Science Research Division). Appropriate amount of protein was used for Dual Luciferase Assay or Western Blot.

Nuclear proteins from CFSC-8B cells, Chang liver cells, H-4-II-E cells and HSC-T6 cells were isolated using Nuclei EZ prep kit (Sigma-Aldrich). Cells were washed with ice cold PBS buffer and then cell dishes were set on ice. One milliliter of ice cold Nuclei EZ lysis buffer was added into each dish. Cells were harvested and lysed by thoroughly scraping each dish with a small bladed cell scraper. Cell lysate from each dish was transferred to a 1.5ml tube and centrifuged at 500g for 5 minutes at 4°C. Supernatant was discarded. Cell pellet was resuspended with 0.5ml of ice cold Nuclei EZ lysis buffer by vortexing briefly at moderate to high speed. Another 1ml of ice cold Nuclei EZ lysis buffer was added into the same tube and the reaction was set on ice for 5 minutes. Cell lysate was centrifuged at 500g for 5 minutes at 4°C and supernatant was discarded. Cell pellet was redissolved in 200µl of ice cold Nuclei EZ storage buffer with 2µl of proteinase inhibitor (Invitrogen). Protein concentration was determined using a protein assay kit (Life Science Research Division). Appropriate amount of nuclear proteins was used for DNA and protein binding reaction.

2.11 Dual Luciferase Assay

Genetic reporter systems are widely used to study eukaryotic gene expression and dual reporters are commonly used to improve experimental accuracy. The term of 'dual reporter' refers to the simultaneous expression and measurement of two individual reporter enzymes within a single tube. The testing reporter is related to some specific experimental conditions while the co-transfected control reporter

provides an internal control that serves as baseline response. In this study, the pRL-TK plasmid (Promega) containing Renilla luciferase with TK promoter was co-transfected with Smad1.pGL2 constructs containing firefly luciferase to normalize variability of transfection efficiency.

Firefly luciferase and Renilla luciferase can generate luminescence using different substrates. Light emission of firefly luciferase is achieved through oxidation of beetle luciferin in a reaction requiring ATP, Mg^{2+} and O_2 while the luminescent reaction catalyzed by Renilla luciferase utilizes coelenterazine, coelenterate-luciferin and O_2 (Fig.13).

This item has been removed due to copyright issues. To view it, refer to its source.

Fig.13. Bioluminescent reactions catalyzed by firefly and *Renilla* luciferases.

Cells in 35mm-dish were transfected with 4 μ g of rat Smad1.pGL2 constructs and 20ng of pRL-TK plasmid. Each transient transfection was performed in triplicates and repeated at least two times. Proteins were isolated 36 hours after transfection and dissolved in 60-100 μ l of lysis buffer, 20-60 μ l of which was used for measurement of luciferase activities using the “Dual-luciferase reporter assay system” (Promega) with a luminometer (ILA911 Luminometer, Tropix Inc, Bedford, MA).

In Dual Luciferase Assay, the activities of firefly and Renilla luciferase were measured sequentially from a single sample. The firefly luciferase activities of rat

with a luminometer (ILA911 Luminometer, Tropix Inc, Bedford, MA). In Dual Luciferase Assay, the activities of firefly and Renilla luciferase were measured sequentially from a single sample. The firefly luciferase activities of rat Smad1.pGL2 constructs were measured first by adding 100 μ l of Luciferase Assay Reagent II (LARII) to generate a stabilized luminescent signal. After quantifying the firefly luminescence, the reaction was stopped and the Renilla luciferase activities were measured by subsequently adding 100 μ l of Stop & Glo® Reagent to the same tube. The Stop & Glo® Reagent produced a stabilized signal from the Renilla luciferase. Therefore, the activity was registered as the ratio of firefly to Renilla luciferase.

2.12 Western blot analysis

In the case of α -Smooth muscle actin (α -SMA), 10 μ g of total proteins isolated from CFSC-8B cells were used in Western blot analysis. Proteins mixed with 4X gel loading buffer (4X = 250mmol/L Tris-HCl, pH=6.8, 8% SDS, 20% glycerol, 0.2% bromophenol blue and 5% β -mercaptoethanol) was boiled for 5 minutes and separated on 12% sodium dodecyl sulfate-polyacrylamide (SDS-polyacrylamide) gel electrophoresis (Mini Trans-blot Electrophoretic Transfer cell, Life Science Research Division) for 1.5-2 hours. The separated protein bands were transferred onto Nitroplus-2000 membrane (Micron Separations Inc. Westborough, MA) by electrophoresis for 1.5 hours at room temperature (Mini Trans-blot Electrophoretic Transfer cell, Life Science Research Division). Nitrocellulose membrane with proteins was blocked by incubating in 5% dry skim milk in TBS-T (Tris-buffered-saline-Tween, 2.42g Tris base, 29.25g NaCl, and 0.5ml Tween 20 per liter) for 1 hour at room temperature. Purpose of blocking is to cover protein

binding sites on nitrocellulose membrane with milk to avoid nonspecific binding of antibody. Subsequently, membranes were incubated O/N at 4°C with primary antibodies against α -SMA at dilution of 1:2000 in 2% skim milk in TBS-T. After washing with TBS-T, membrane was incubated with secondary antibody sheep anti-mouse IgG at dilution of 1:1000 for 1 hour at room temperature (antibodies were purchased from Santa Cruz Biotechnology, Inc). The protein-antibody complexes were detected by enhanced chemiluminescence (ECL system, GE health care) and the densitometry of the bands was determined using ImageJ program.

2.13 DNA Labeling and Purification

DNA fragments to be labeled were either generated by PCR, or purchased from Invitrogen. The commercial DNA fragments were ordered as complementary pairs. To generate double-stranded DNA, two complementary oligos were first heated at 95°C for 5 minutes in a heating block (Isotemp*125D heating block, Fisher Scientific Limited, Nepean, ON), and then the complementary oligos were annealed in the heating block (turning power off) by slowly cooling down to room temperature.

Double-stranded DNA fragments were labeled at their 5'-ends using T4 polynucleotide kinase (Promega), which can catalyze the transfer of γ -phosphate of ATP to the 5'-terminus of DNA or RNA molecules. The labeling reaction was performed in a volume of 10 μ l containing 100ng double-stranded DNA, 4 μ l of γ -³²P ATP (250 μ Ci, GE health care), 1 μ l of T4 polynucleotide kinase, 1 μ l of 10X T4 polynucleotide kinase buffer, and sterile distilled water. Labeling reaction was performed at 37°C for 1 hour and stopped by adding 2 μ l of 0.5M EDTA.. Another 90 μ l of TE buffer was added into reaction to dilute DNA into 1ng/ μ l.

Labeled DNA was purified by passing through Microspin G-25 column (GE health

care) according to the manufacturer's instructions. The labeling quality was checked by a Model 3A Geiger counter (Measurements, Inc. Sweetwater, Texas). Well labeled probes should register between 1 and 2 kcps.

2.14 Electrophoretic Mobility Shift Assay (EMSA)

DNA and protein binding reaction was performed at room temperature for 20 minutes. Each reaction contained 1ng of labeled DNA probe, 5µg of nuclear extracts, 1µg of poly (dI-dC) (GE health care), 6mM MgCl₂ and appropriate amount of 1X binding buffer (12mM Hepes pH=8.0, 60mM KCl, 2.5mM EDTA, 14% glycerol, 2mM dTT, and 0.5mM PMSF) to make final volume at 20µl. Poly (dI-dC) was used as a nonspecific competitor to minimize the binding of nonspecific proteins to the labeled DNA probe.

For competition assay, 100- or 250-fold molar excess of cold probes, unlabeled consensus sequences of putative regulatory factors were pre-incubated with nuclear extracts at room temperature for 5 minutes, and then labeled probes were added into the reaction for another 15 minutes.

EMSA was carried out using 6% polyacrylamide gel in 0.5X TBE buffer. Before loading samples, gel was pre-run at 100v for 30 minutes. Samples mixed with loading dye were carefully loaded. After 2-3 hours of electrophoresis at 200v at 4°C, the gel was dried for 1-2 hours by Slab Gel Dryer-2000 (Fisher Scientific Limited) at 50 °C and exposed to X-ray film (Kodak Biomax XAR film) for 1-3 days.

2.15 Sequence Analysis

Genome sequence information and gene information were obtained from Ensembl and GenBank. Prediction of transcriptional factor binding sites was studied

using the on-line tool in TFSEARCH (Table3).

Table3. Websites of sequence analysis

Name	Website
Ensembl	http://www.ensembl.org/index.html
GenBank	http://www.ncbi.nlm.nih.gov/Genbank/index.html
TFSEARCH	http://www.cbrc.jp/research/db/TFSEARCH.html

III. RESULTS

1. Rat genomic DNA isolation

Rat genomic DNA isolated from kidney tissue was analyzed by Agarose Gel Electrophoresis and restriction enzyme digestion (Fig.14). The kidney from an adult Spragur-Dawley rat was kindly provided by Ms. Jianghong Fan (University of Manitoba, Canada). Approximately 200ng of uncut genomic DNA was run on a 0.8% agarose gel. The single band near the well was characteristic of whole uncut genomic DNA and the lane showed slight smearing which indicate most of the DNA was unbroken. The intense band on the bottom of the gel indicated RNA. Genomic DNA was analyzed by restriction enzyme digestion as well. The reaction was performed with 200ng of DNA and 1 μ l of 10u/ μ l EcoRI at 37°C for 1 hour and subsequently run on a 0.8% agarose gel. The genomic DNA was cut into several fragments, some of which could be observed in this gel. All these features of the uncut and cut genomic DNA indicated that the preparations were good quality.

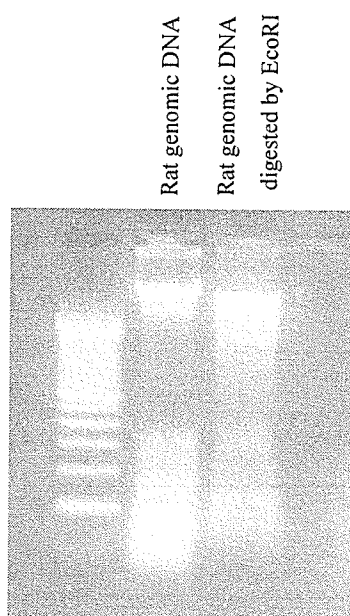


Fig.14. Rat genomic DNA analysis. Genomic DNA was analyzed on a 0.8% agarose gel. Uncut DNA showed single band near the well and light smearing. EcoRI digested DNA showed several fragments. These features indicated the quality of preparations was good.

2. Cloning of rat Smad1 5'-flanking region and intron1 region

2.1 PCR cloning

Rat genomic DNA sequence containing Smad1 gene locus was obtained from Ensembl website (Ensembl Gene ID, ENSRNOG00000018483). According to this sequence, twelve pairs of PCR primers were designed using Oligo5 program to clone rat Smad1 5'-flanking region and intron1 region. Rat genomic DNA isolated from kidney tissue was used as the PCR template.

Primer sequences and PCR conditions are shown in Table4. All the primers contained restriction enzyme sites at their 5'-ends which were shown by underlining. Primer design and the organization of the PCR cloning were shown in Fig.15. The first nucleotide acids at 5'-flanking region/exon1 boundary and intron1/exon2 boundary were set as F+1 and I+1 respectively.

Primers for cloning of the 5'-flanking region contained XhoI and HindIII sites. A series of 5'-end deletional PCR fragments covered F-1005 to +93, F-605 to +93, F-430 to +93, F-163 to +93 and F-56 to +93 were generated by different upper primers F1-F5 with XhoI sites and same lower primer F0 with HindIII site. Primers for intron1 cloning contained BglII and HindIII sites. A series of 5'-end truncated PCR fragments covered I-2015 to +59, I-1109 to +59, I-736 to +59, I-610 to +59, I-490 to +59, I-389 to +59, and I-273 to +59 were amplified by different upper primers I1-I7 with BglII sites and same lower primer I0 with HindIII site. PCR results are shown in Fig.16 and 17.

Table4. Primer sequences and PCR conditions of cloning of rat Smad1 5'-flanking and intron1 region

Name	Primer Sequences (5'-3')	Location	Temp.	Product
5-flanking region				
F0	CGC GCA AGC TTC CGT TCC CAG TCA CC	F+93		
F1	CGG GCC TCG CAA GCT GGG CGG TTC	F-56	66°C	171bp
F2	CGC GGC TCG AGA ATT TGG AGA GGA T	F-163	58°C	277bp
F3	CAT ACC TCG AGC TTT TGT CAT TTG CCG C	F-430	64°C	544bp
F4	CCT GCC TCG AGT GCA CAA GGC CAG GT	F-605	66°C	720bp
F5	TCT TTC TCG AGG TGG GGC ATT TAC TCC T	F-1005	64°C	1120bp
Intron1 region				
I0	CTT CTA AGC TTT CGG TTC GGA AAG GTA GG	I+59		
I1	ACT TAA GAT CTC CAA GCC CAA ATA CCA GTC	I-273	57°C	352bp
I2	CTT TGA GAT CTG CTT GGG TGC TGT TGG	I-389	58°C	470bp
I3	GTG CTA GAT CTC TGT TGG GGC TTG TCA CTT	I-490	58°C	570bp
I4	TGC CGA GAT CTT TGG TCG CAT TTT TAC GC	I-610	58°C	691bp
I5	GCT CAA GAT CTG CTC TGA GAA CGG GAC G	I-736	59°C	819bp
I6	TTA TGA GAT CTT TCG GGC TGG GGA TT	I-1109	59°C	1189bp
I7	GGT ACA GAT CTC ACC CCA ATA CAC TCA TTT CA	I-2015	58°C	2096bp

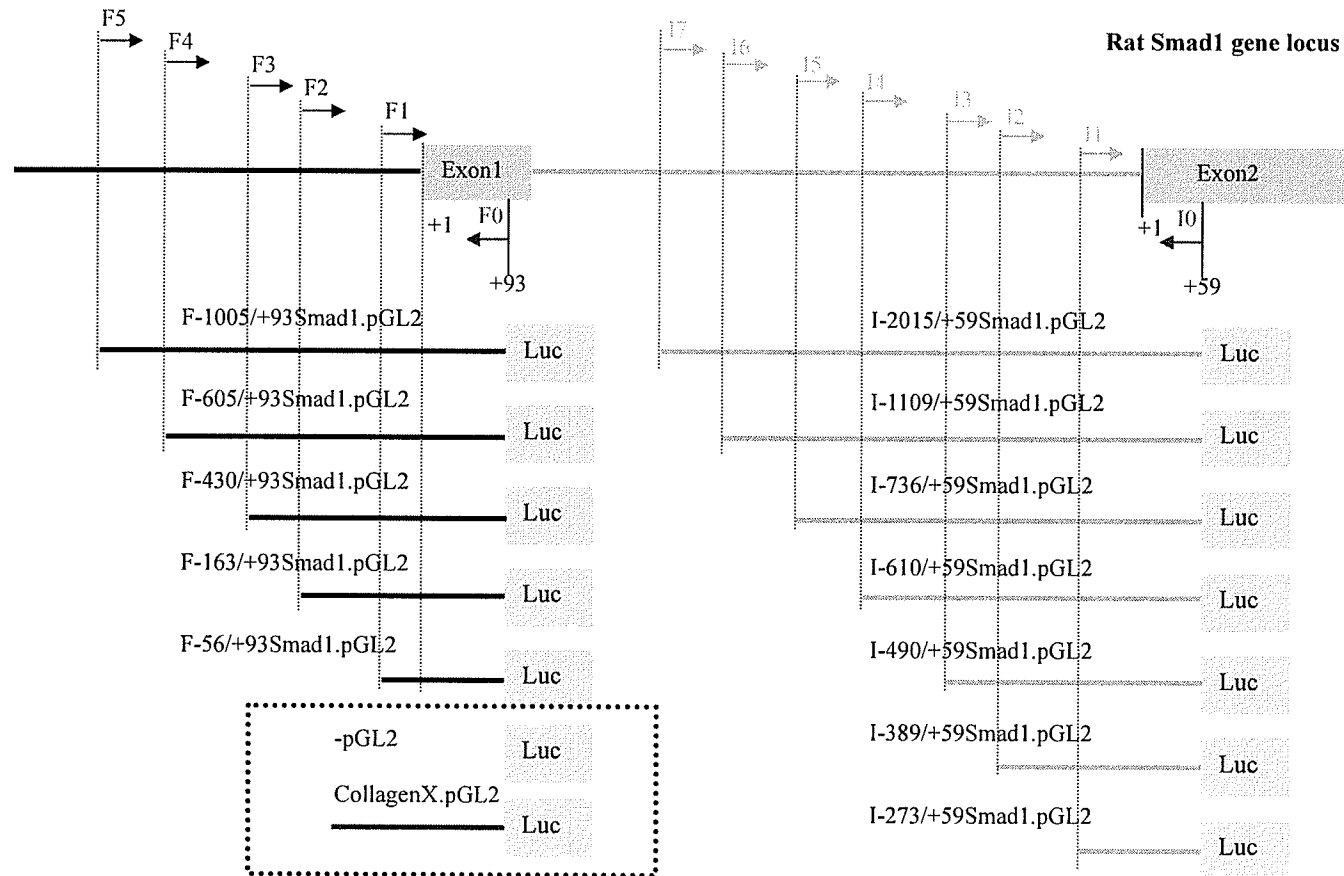


Fig.15. Schematic organization of cloning of rat Smad1 5'-flanking region and intron1 region.

(1) 5'-flanking region (blue line), intron1 region (green line) and two exons (orange boxes) of rat Smad1 gene were showing in the figure. The first nucleotide acids at 5'-flanking region/exon1 boundary and intron1/exon2 boundary were set as F+1 and I+1 respectively. Sequence upstream was minus and downstream was plus. (2) Five pairs of PCR primers (F1-F5, blue arrows; F0, red arrow) were designed from 5'-flanking region and seven pairs (I1-I7, green arrows; I0, red arrow) were designed from intron1 region. (3) A series of 5'-end truncated PCR products (blue and green lines) were transferred into pGL2-Basic vector (yellow boxes showing "Luc") to generate Smad1.pGL2 constructs. The first letter of the construct name indicated the region where the product was from. F- represented 5'-flanking region while I- represented intron1 region. The numbers adjacent to the letter explained the area that individual PCR fragment covered. For the fragments from each region, they had same 3'-end (F+93 or I+59) but different 5'-ends. (4) -pGL2 and CollagenX.pGL2 were the negative and positive control used in Dual Luciferase Assay. -pGL2 was pGL2-Basic vector without any transcriptional activities while CollagenX.pGL2 was generated by fusing CollagenX promoter region (purple line) and pGL2-Basic vector.

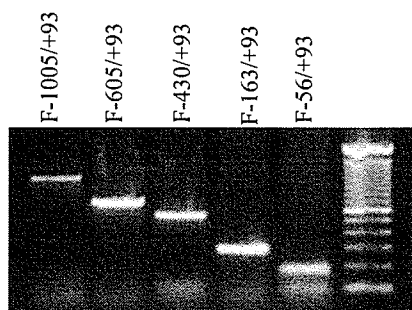


Fig.16. PCR cloning of rat Smad1 5'-flanking region. Five deletional fragments from 5'-flanking region were cloned using PCR method, covering F-1005 to +93, F-605 to +93, F-430 to +93, F-163 to +93 and F-56 to +93.

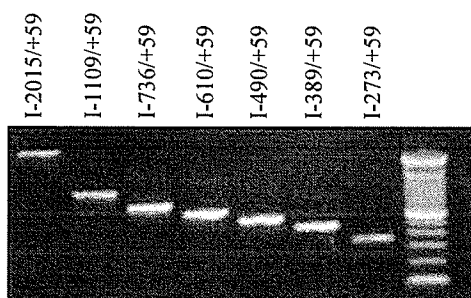


Fig.17. Seven deletional fragments from intron1 region were cloned using PCR method, covering I-2015 to +59, I-1109 to +59, I-736 to +59, I-610 to +59, I-490 to +59, I-389 to +59, and I-273 to +59.

2.2 Subcloning and reconstruction of plasmid

The pGL2-Basic vector (Fig.18) containing the firefly luciferase reporter gene (Promega) was used to generate rat Smad1.pGL2 constructs.

PCR products of rat Smad1 5'-flanking region and pGL2-Basic vector were digested with XhoI and HindIII (Promega) to create the cohesive ends. Subsequently, both DNA preparations with complementary sticky ends were incorporated into Smad1.pGL2 constructs, including F-1005/+93Smad1.pGL2, F-605/+93Smad1.pGL2, F-430/+93Smad1.pGL2, F-163/+93Smad1.pGL2, and F-56/+93Smad1.pGL2.

PCR products of rat Smad1 intron1 region were first subcloned into pCR4-Topo vector (Fig.19). pCR4-Topo vector has a 3'-T overhangs that can directly ligate Taq-amplified PCR products. Subsequently, the corresponding fragments were excised from the pCR4-Topo vector and ligated into BglII/HindIII site of the pGL2-Basic vector. Seven different constructs of rat Smad1 intron1 region were generated: I-2015/+59Smad1.pGL2, I-1109/+59Smad1.pGL2, I-736/+59Smad1.pGL2, I-610/+59Smad1.pGL2, I-490/+59Smad1.pGL2, I-389/+59Smad1.pGL2, and I-273/+59Smad1.pGL2.

Positive constructs were selected by restriction enzyme digestion (Fig.20-22) and confirmed by DNA sequencing (University of Calgary, Fig.23). GLprimer1 (5'-TGT ATC TTA TGG TAC TGT AAC TG-3') and GLprimer2 (5'-CTT TAT GTT TTT GGC GTC TTC CA-3') from Promega were used as sequencing primers.

**This item has been removed
due to copyright issues. To
view it, refer to its source.**

Fig.18. pGL2-Basic vector map.

**This item has been removed
due to copyright issues. To
view it, refer to its source.**

Fig.19. pCP4-Topo vector map.

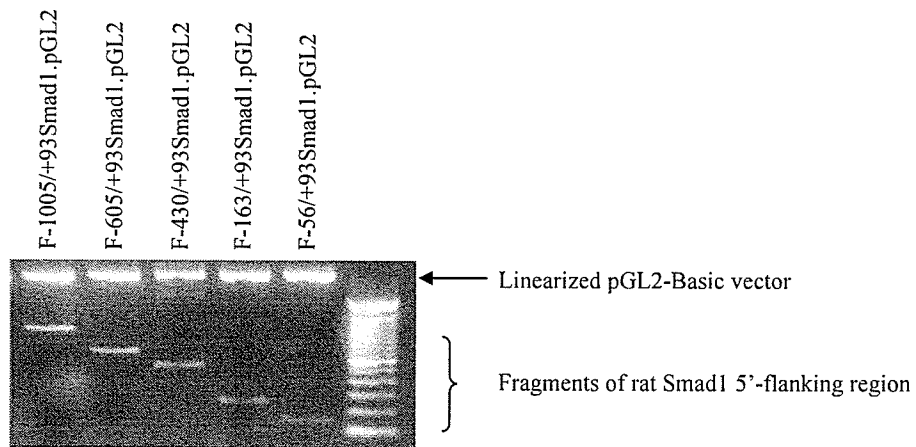


Fig.20. Smad1.pGL2 constructs containing fragments of rat Smad1 5'-flanking region. Constructs were digested with XhoI and HindIII. Bands with the same size on the top were linearized pGL2-Basic vectors. Lower bands were different in length representing the insert fragments from rat Smad1 5'-flanking region.

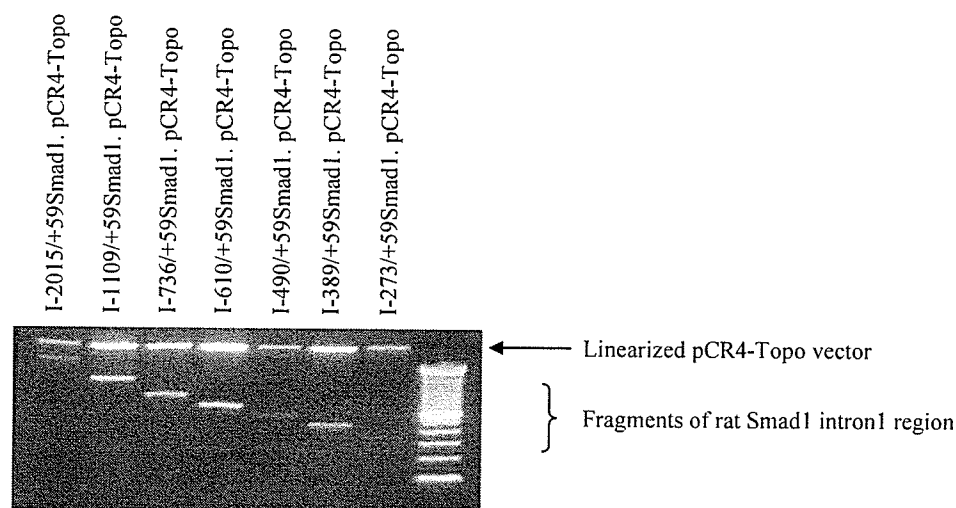


Fig.21. Smad1.pCR4-Topo constructs containing fragments of rat Smad1 intron1 region. Constructs were digested with EcoRI. Bands with the same size on the top were linearized pCR4-Topo vectors. Bands on the bottom were the insert fragments from rat Smad1 intron1 region.

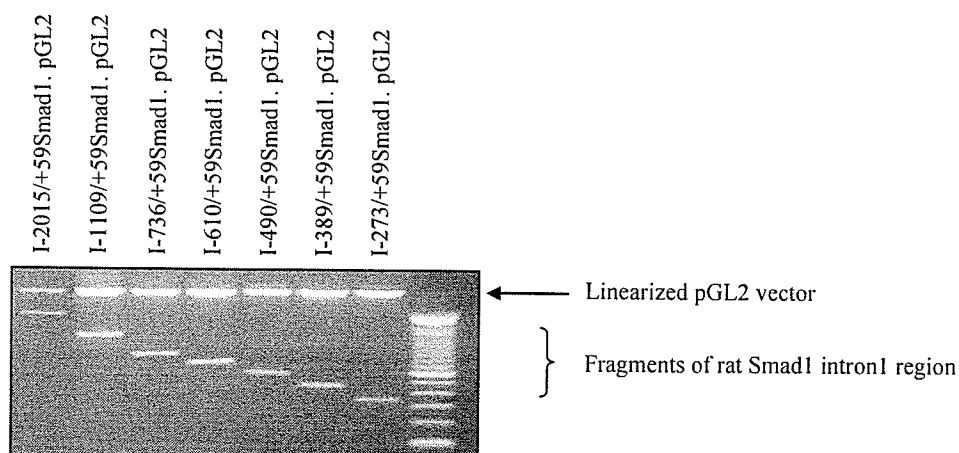


Fig.22. Smad1.pGL2 constructs containing fragments of rat Smad1 intron1 region. Constructs were digested with BglII and HindIII. Bands on the top were linearized pGL2-Basic vectors. Bands on the bottom were the insert fragments from rat Smad1 intron1 region.

F-56/+93Smad1.pGL2

Query	1145	GGCTGGGCGGTTCCGACTGCCTGACGCCAGCGCGACCAGATCAATCCAGGCTCGGGGAGC	1204
Sbjct	42	GGCTGGGCGGTTCCGACTGCCTGACGCCAGCGCGACCAGATCAATCCAGGCTCGGGGAGC	101
Query	1205	GAGCGGGCGCACCAAGGCGAGGCCAGGGCCGAGGCGCGGGGACGGCGCCGAGAGCTAAG	1264
Sbjct	102	GAGCGGGCGCACCAAGGCGAGGCCAGGGCCGAGGCGCGGGGACGGCGCCGAGAGCTAAG	161
Query	1265	CGAAGCGCTGGGACGGTGACTGGGAAC	1291
Sbjct	162	CGAAGCGCTGGGACGGTGACTGGGAAC	188

F-163/+93Smad1.pGL2

Query	1038	GAATTTGAGAGGATCCCTAAGCAGCAGCGGGCGCGCGGTCGCGTGTGAGCGAGGGC	1097
Sbjct	43	GAATTTGAGAGGATCCCTAAGCAGCAGCGGGCGCGCGGTCGCGTGTGAGCGAGGGC	102
Query	1098	GAGCGGCGGGCGGGTTGTGCGCCGGTATTGGCAGGCGGGCTGTGGAGGCTGGGCGTTTC	1157
Sbjct	103	GAGCGGCGGGCGGGTTGTGCGCCGGTATTGGCAGGCGGGCTGTGGAGGCTGGGCGTTTC	162
Query	1158	CGACTGCCTGACGCCAGCGCGACCAGATCAATCCAGGCTCGGGGAGCGAGCGGGCGCACC	1217
Sbjct	163	CGACTGCCTGACGCCAGCGCGACCAGATCAATCCAGGCTCGGGGAGCGAGCGGGCGCACC	222
Query	1218	AAGGCGAGGCCAGGGCCGAGGCGCGGGGACGGCGGCCGAGAGCTAAGCGAAGCGCTGGGA	1277
Sbjct	223	AAGGCGAGGCCAGGGCCGAGGCGCGGGGACGGCGGCCGAGAGCTAAGCGAAGCGCTGGGA	282
Query	1278	CGGTGACTGGGAACGGA	1294
Sbjct	283	CGGTGACTGGGAACGGA	299

F-430/+93Smad1.pGL2

Query	771	GCTTTTGTCAATTGCCCGAGCAAGGGTACATACTAGAAATAAATGCATTTACCTAGCAGT	830
Sbjct	44	GCTTTTGTCAATTGCCCGAGCAAGGGTACATACTAGAAATAAATGCATTTACCTAGCAGT	103
Query	831	TAACAGGTCACCCAGTTTGGGGCTCAGGCCACATGGAGGGCCGGTAAATCACTGGGTCAG	890
Sbjct	104	TAACAGGTCACCCAGTTTGGGGCTCAGGCCACATGGAGGGCCGGTAAATCACTGGGTCAG	163
Query	891	TGATTGCAGGGCCATTAAATGGGTTTTAGGACGATTGGGTGGGTTTGATGGTCGTGTT	950
Sbjct	164	TGATTGCAGGGCCATTAAATGGGTTTTAGGACGATTGGGTGGGTTTGATGGTCGTGTT	223
Query	951	TGAGAACATTTCAGAAAATAACACAGGCTTTGGAACCTTGACTGGCAGCTAGGCAGGGACT	1010

Sbjct	224	TGAGAACATTTCAGAAAATAACACAGGCTTTGGAACCTGACTGGCAGCTAGGCAGGGACT	283	
Query	1011	ACAGAGTCAGAGGCAGGCGGGCCGGGAATTTGGAGAGGATCCCTAAGCAGCAGCGGGC	1070	
Sbjct	284	ACAGAGTCAGAGGCAGGCGGGCCGGGAATTTGGAGAGGATCCCTAAGCAGCAGCGGGC	343	
Query	1071	GCGCGCGGTGCGCGTGTGAGCGAGGGCGAGCGGGCGGGTGTGCGCCGGGTATTGGC	1130	
Sbjct	344	GCGCGCGGTGCGCGTGTGAGCGAGGGCGAGCGGGCGGGTGTGCGCCGGGTATTGGC	403	
Query	1131	AGGCGGGCTGTGGAGGCTGGGCGGTTCCGACTGCCTGACGCCAGCGCAGCAGATCAATC	1190	
Sbjct	404	AGGCGGGCTGTGGAGGCTGGGCGGTTCCGACTGCCTGACGCCAGCGCAGCAGATCAATC	463	
Query	1191	CAGGCTCGGGGAGCGAGCGGGCGCACCAAGGCGAGGCCAGGGCCAGGCGCGGGGACGGC	1250	
Sbjct	464	CAGGCTCGGGGAGCGAGCGGGCGCACCAAGGCGAGGCCAGGGCCAGGCGCGGGGACGGC	523	
Query	1251	GGCCGAGAGCTAAGCGAAGCGCTGGGACGGTGACTGGGAACG	1292	
Sbjct	524	GGCCGAGAGCTAAGCGAAGCGCTGGGACGGTGACTGGGAACG	565	

F-605/+93Smad1.pGL2

Query	596	TGCACAAGGCCAGGTTTCGATTCCCAGAATCTTATAAATAAAAAATATGTGTAGATCTATAA	655	
Sbjct	41	TGCACAAGGCCAGGTTTCGATTCCCAGAATCTTATAAATAAAAAATATGTGTAGATCTATAA	100	
Query	656	TCCCAGATCTCACGATGTGGAGGCTGGAGGAAGAGAAGTTCAAGGTCTTTTCCCCTTCAA	715	
Sbjct	101	TCCCAGATCTCACGATGTGGAGGCTGGAGGAAGAGAAGTTCAAGGTCTTTTCCC-TTCAA	159	
Query	716	AGTGAGTTCGGGGTCATCCTGGGCTACAAGAGACCATGTATCAAACATACAAATGGCTTT	775	
Sbjct	160	AGTGAGTTCGGGGTCATCCTGGGCTACAAGAGACCATGTATCAAACATACAAATGGCTTT	219	
Query	776	TGTCATTTGCCGCAGCAAGGGTACATACTAGAAAATAAATGCATTTACCTAGCAGTTAACA	835	
Sbjct	220	TGTCATTTGCCGCAGCAAGGGTACATACTAGAAAATAAATGCATTTACCTAGCAGTTAACA	279	
Query	836	GGTCACCCAGTTTGGGGCTCAGGCCACATGGAGGGCCGGTAAATCACTGGGTCAGTGATT	895	
Sbjct	280	GGTCACCCAGTTTGGGGCTCAGGCCACATGGAGGGCCGGTAAATCACTGGGTCAGTGATT	339	
Query	896	GCAGGGCCATTAAATGGGTTTTAGGACGATTGGGTGGGTTTGATGGTCGTGTTGAGA	955	
Sbjct	340	GCAGGGCCATTAAATGGGTTTTAGGACGATTGGGTGGGTTTGATGGTCGTGTTGAGA	399	
Query	956	ACATTTTCAGAAAATAACACAGGCTTTGGAACCTGACTGGCAGCTAGGCAGGGACTACAGA	1015	
Sbjct	400	ACATTTTCAGAAAATAACACAGGCTTTGGAACCTGACTGGCAGCTAGGCAGGGACTACAGA	459	

Query	1016	GTCAGAGGCAGGCGCGGGCCGGAATTTGGAGAGGATCCCTAAGCAGCAGCGGGCGCGCG	1075
Sbjct	460	GTCAGAGGCAGGCGCGGGCCGGAATTTGGAGAGGATCCCTAAGCAGCAGCGGGCGCGCG	519
Query	1076	CGGTGCGCGTGTGAGCGAGGGCGAGCGCGGGCGGGTTGTGCGCCGGGTATTGGCAGGCG	1135
Sbjct	520	CGGTGCGCGTGTGAGCGAGGGCGAGCGCGGGCGGGTTGTGCGCCGGGTATTGGCAGGCG	579
Query	1136	GGCTGTGGAGGCTGGGCGGTTCCGACTGCCTGACGCCAGCGCGACCAGATCAATCCAGGC	1195
Sbjct	580	GGCTGTGGAGGCTGGGCGGTTCCGACTGCCTGACGCCAGCGCGACCAGATCAATCCAGGC	639
Query	1196	TCGGGGAGCGAGCGGGCGCACCAAGGCGAGGCCAGGGCCGAGGCGCGGGGACGGCGGCCG	1255
Sbjct	640	TCGGGGAGCGAGCGGGCGCACCAAGGCGAGGCCAGGGCCGAGGCGCGGGGACNGCGGCCG	699
Query	1256	AGAGCTAAGCGAAGCGCTGGGACGGTGACTGGGAACGGA	1294
Sbjct	700	AGAGCTAAGCGAAGCGCTGGGACGGTGACTGGGAACGGA	738

F-605/+93Smad1.pGL2

Query	196	GTGGGGCATTACTCCTAGAATAGTTCTTCTCTGAACATAACTCAACATTATGATAAGTG	255
Sbjct	41	GTGGGGCATTACTCCTAGAATAGTTCTTCTCTGAACATAACTCAACATTATGATAAGTG	100
Query	256	TCCAAGATTTTCATGCCAACTTCCAGAACTGACCAAAAAAAGTTGATTGTTTCAAGAG	315
Sbjct	101	TCCAAGATTTTCATGCCAACTTCCAGAACTGACCAAAAAAAGTTGATTGTTTCAAGAG	160
Query	316	GGTGTTTTGGTGTCTGAACAGATCAGTAAGATGTGAGCTCCTTTGTGTTGTGTATTTT	375
Sbjct	161	GGTGTTTTGGTGTCTGAACAGATCAGTAAGATGTGAGCTCCTTTGTGTTGTGTATTTT	220
Query	376	TGTTTTGTGGGGTTTTTGTGTTGGGGGTGGGGGTGGAACAGGGGCAAGACCCTGTCATC	435
Sbjct	221	TGTTTTGTGGGGTTTTTGTGTTGGGGGTGGGGGTGGAACAGGGGCAAGACCCTGTCATC	280
Query	436	TAAAAATTTACACCAAAATTTATAGTACTAAAAGGGTATTTCCCAATTCAAGGCCAGCC	495
Sbjct	281	TAAAAATTTACACCAAAATTTATAGTACTAAAAGGGTATTTCCCAATTCAAGGCCAGCC	340
Query	496	TGGGCTAAAGGGAGAGATCCCGTTCCAAGCCGGCAAGCAAGTGAACAAAAGGGCTA	555
Sbjct	341	TGGGCTAAAGGGAGAGATCCCGTTCCAAGCCGGCAAGCAAGTGAACAAAAGGGCTA	400
Query	556	GGTGGTGAAGCTCAGTTGTAAAGTTCTTCTGCCCAGGATGCACAAGGCCAGGTTTCGAT	615
Sbjct	401	GGTGGTGAAGCTCAGTTGTAAAGTTCTTCTGCCCAGGATGCACAAGGCCAGGTTTCGAT	460
Query	616	TCCCAGAATCTTATAAATAAAAAATATGTGTAGATCTATAATCCCAGATCTCACGATGTGG	675
Sbjct	461	TCCCAGAATCTTATAAATAAAAAATATGTGTAGATCTATAATCCCAGATCTCACGATGTGG	520

Query	676	AGGCTGGAGGAAGAGAAGTTCAAGGTCTTTT---CCCCTTCAAAGTGAGTTCGGGGTCATC	733
Sbjct	521	AGGCTGGAGGAAGAGAAGTTCAAGGTCTTTTCCCCCTTCNAAGTGAGTTCGGGGTCATC	580
Query	734	CTGGGCTA-CAAGAGACC-ATGTATCAAACATACAAATGGC-TTTTGTCAATTGCCGCAG	790
Sbjct	581	CTGGGCTACCAAGAGACCNATGTATCAAACATACAAATGGCTTTTGTCAATTGCCGCAG	640
Query	791	CAA-GGGTACATACTAGAAATAAATGC-ATTTACCTAGCAGTTAA	833
Sbjct	641	CAAGGGGTACATACTAGAAATAAATGCAATTTACCTAGCAGNTAA	685

I-273/+59Smad1.pGL2

Query	1728	CTCCAAGCCCAAATACCAGTCGGTCGCAGTGGGTGCTCTGCCTTTAACTGGTGAACCTTTA	1787
Sbjct	41	CTCCAAGCCCAAATACCAGTCGGTCGCAGTGGGTGCTCTGCCTTTAACTGGTGAACCTTTA	100
Query	1788	GGAGAACGTAGCTCAGTTCCTGGTGTACAGGACACAAATGTTTCATGAACCTCAGCATGGG	1847
Sbjct	101	GGAGAACGTAGCTCAGTTCCTGGTGTACAGGACACAAATGTTTCATGAACCTCAGCATGGG	160
Query	1848	GTTGCTCTCTAATGAAGTCACCGAATTGGTCTAAGCCAAGCGTGGACCTCTACAGGGTCA	1907
Sbjct	161	GTTGCTCTCTAATGAAGTCACCGAATTGGTCTAAGCCAAGCGTGGACCTCTACAGGGTCA	220
Query	1908	GACTTAAGTAGTTAGTACCTGGCAGCGACTAATTAGCTAATTGTGACGCAGTTAGAGTAA	1967
Sbjct	221	GACTTAAGTAGTTAGTACCTGGCAGCGACTAATTAGCTAATTGTGACGCAGTTAGAGTAA	280
Query	1968	TTTTAAATCAGCTTCTTGCTTTTCTTTTCGCAGGTGCTAACTGGGTCTTCCCTGAAGCAG	2027
Sbjct	281	TTTTAAATCAGCTTCTTGCTTTTCTTTTCGCAGGTGCTAACTGGGTCTTCCCTGAAGCAG	340
Query	2028	AAGGAACAGTATTTCTACCTTTCCGAACCGAA	2060
Sbjct	341	AAGGAACAGTATTTCTACCTTTCCGAACCGAA	373

I-389/+59Smad1.pGL2

Query	1612	GCTTGGGTGCTGTTGGGAGCAATCTATTCCCTTCCTCTTGTAAGCTGCCAATCCTGTTG	1671
Sbjct	46	GCTTGGGTGCTGTTGGGAGCAATCTATTCCCTTCCTCTTGTAAGCTGCCAATCCTGTTG	105
Query	1672	CCTCAAGCCCCACCCACCTTATAATCTAACTTAACCCTAGTGAGTTACTTAGAGGCTCC	1731
Sbjct	106	CCTCAAGCCCCACCCACCTTATAATCTAACTTAACCCTAGTGAGTTACTTAGAGGCTCC	165
Query	1732	AAGCCCAAATACCAGTCGGTCGCAGTGGGTGCTCTGCCTTTAACTGGTGAACCTTAGGAG	1791
Sbjct	166	AAGCCCAAATACCAGTCGGTCGCAGTGGGTGCTCTGCCTTTAACTGGTGAACCTTAGGAG	225
Query	1792	AACGTAGCTCAGTTCCTGGTGTACAGGACACAAATGTTTCATGAACCTCAGCATGGGGTTG	1851

Sbjct	226	 AACGTAGCTCAGTTCCTGGTGTACAGGACACAAATGTTTCATGAACTCAGCATGGGGTTG	285
Query	1852	CTCTCTAATGAAGTCACCGAATTGGTCTAAGCCAAGCGTGGACCTCTACAGGGTCAGACT	1911
Sbjct	286	 CTCTCTAATGAAGTCACCGAATTGGTCTAAGCCAAGCGTGGACCTCTACAGGGTCAGACT	345
Query	1912	TAAGTAGTTAGTACCTGGCAGCCTAATTAGCTAATTGTGACGCAGTTAGAGTAATTTT	1971
Sbjct	346	 TAAGTAGTTAGTACCTGGCAGCCTAATTAGCTAATTGTGACGCAGTTAGAGTAATTTT	405
Query	1972	AAATCAGCTTCTTGCTTTTCTTTCGAGGTGCTAACTGGGTCTTCCCTGAAGCAGAAGG	2031
Sbjct	406	 AAATCAGCTTCTTGCTTTTCTTTCGAGGTGCTAACTGGGTCTTCCCTGAAGCAGAAGG	465
Query	2032	AACAGTATTCCTACCTTTCCGAACCGAA	2060
Sbjct	466	 AACAGTATTCCTACCTTTCCGAACCGAA	494

I-490/+59Smad1.pGL2

Query	1511	TCTGTTGGGGCTTGTCACCTTCGTTTACATATGAGGTGGTAATATGCAAATCAGCCAAGT	1570
Sbjct	40	 TCTGTTGGGGCTTGTCACCTTCGTTTACATATGAGGTGGTAATATGCAAATCAGCCAAGT	99
Query	1571	GGATTTCTGTGCCCCGAACACTCGCATGACCTTTGTCCAGGGCTTGGGTGCTGTTGGGAG	1630
Sbjct	100	 GGATTTCTGTGCCCCGAACACTCGCATGACCTTTGTCCAGGGCTTGGGTGCTGTTGGGAG	159
Query	1631	CAATCTATTCCTTCCTCTTGTAAGCTGCCAATCCTGTTGCCTCAAGCCCCACCCACCC	1690
Sbjct	160	 CAATCTATTCCTTCCTCTTGTAAGCTGCCAATCCTGTTGCCTCAAGCCCCACCCACCC	219
Query	1691	TTATAATCTAACTTAACCCTAGTGAGTTACTTAGAGGCTCCAAGCCCAAATACCAGTCGG	1750
Sbjct	220	 TTATAATCTAACTTAACCCTAGTGAGTTACTTAGAGGCTCCAAGCCCAAATACCAGTCGG	279
Query	1751	TCGCAGTGGGTGCTCTGCCTTTAACTGGTGAACCTTAGGAGAACGTAGCTCAGTTCCTGG	1810
Sbjct	280	 TCGCAGTGGGTGCTCTGCCTTTAACTGGTGAACCTTAGGAGAACGTAGCTCAGTTCCTGG	339
Query	1811	TGTCACAGGACACAAATGTTTCATGAACTCAGCATGGGGTTGCTCTCTAATGAAGTCACCG	1870
Sbjct	340	 TGTCACAGGACACAAATGTTTCATGAACTCAGCATGGGGTTGCTCTCTAATGAAGTCACCG	399
Query	1871	AATTGGTCTAAGCCAAGCGTGGACCTCTACAGGGTCAGACTTAAGTAGTTAGTACCTGGC	1930
Sbjct	400	 AATTGGTCTAAGCCAAGCGTGGACCTCTACAGGGTCAGACTTAAGTAGTTAGTACCTGGC	459
Query	1931	ACGCACTAATTAGCTAATTGTGACGCAGTTAGAGTAATTTAAATCAGCTTCTTGCTTTT	1990

Sbjct	460	ACGCACTAATTAGCTAATTGTGACGCAGTTAGAGTAATTTTAAATCAGCTTCTTGTCTTT	519
Query	1991	TCTTTCGCAGGTGCTAACTGGGTCTTCCCTGAAGCAGAAGGAACAGTATTTCTACCTTT	2050
Sbjct	520	TCTTTCGCAGGTGCTAACTGGGTCTTCCCTGAAGCAGAAGGAACAGTATTTCTACCTTT	579
Query	2051	CCGAACCGAA	2060
Sbjct	580	CCGAACCGAA	589

I-610/+59Smad1.pGL2

Query	1390	TTTGGTCGCATTTTACGCTGTAGTCTTCTGTACCTGTCGGCTCCTGTGTATCAGTCCAT	1449
Sbjct	51	TTTGGTCGCATTTT-CGCTGTAGTCTTCTGTACCTGTCGGCTCCTGTGTATCAGTCCAT	109
Query	1450	CTCCCTCCAGGAGTGTGCGTTGCTGGAGTGGTTAGCTTCTTTGCAGTTTGTGCTTGGT	1509
Sbjct	110	CTCCCTCCAGGAGTGTGCGTTGCTGGAGTGGTTAGCTTCTTTGCAGTTTGTGCTTGGT	169
Query	1510	ATCTGTTGGGGCTTGTCACTTTCGTTTACATATGAGGTGGTAATATGCAAATCAGCCAAG	1569
Sbjct	170	ATCTGTTGGGGCTTGTCACTTTCGTTTACATATGAGGTGGTAATATGCAAATCAGCCAAG	229
Query	1570	TGGATTTCTGTGCCCCGAAACACTCGCATGACCTTTGTCCAGGGCTTGGGTGCTGTTGGGA	1629
Sbjct	230	TGGATTTCTGTGCCCCGAAACACTCGCATGACCTTTGTCCAGGGCTTGGGTGCTGTTGGGA	289
Query	1630	GCAATCTATTCCCTTCCTCTTGTAAGCTGCCAATCCTGTTGCCTCAAGCCCCACCCACC	1689
Sbjct	290	GCAATCTATTCCCTTCCTCTTGTAAGCTGCCAATCCTGTTGCCTCAAGCCCCACCCACC	349
Query	1690	CTTATAATCTAACTTAACCCTAGTGAGTACTTAGAGGCTCCAAGCCCAAATACCAGTCG	1749
Sbjct	350	CTTATAATCTAACTTAACCCTAGTGAGTACTTAGAGGCTCCAAGCCCAAATACCAGTCG	409
Query	1750	GTCGCAGTGGGTGCTCTGCCTTTAACTGGTGAACCTTAGGAGAACGTAGCTCAGTTCCTG	1809
Sbjct	410	GTCGCAGTGGGTGCTCTGCCTTTAACTGGTGAACCTTAGGAGAACGTAGCTCAGTTCCTG	469
Query	1810	GTGTCACAGGACACAAATGTTTCATGAACTCAGCATGGGGTTGCTCTCTAATGAAGTCACC	1869
Sbjct	470	GTGTCACAGGACACAAATGTTTCATGAACTCAACATGGGGTTGCTCTCTAATGAAGTCACC	529
Query	1870	GAATTGGTCTAAGCCAAGCGTGGACCTCTACAGGGTCAGACTTAAGTAGTTAGTACCTGG	1929
Sbjct	530	GAATTGGTCTAAGCCAAGCGTGGACCTCTACAGGGTCAGACTTAAGTAGTTAGTACCTGG	589
Query	1930	CACGCACTAATTAGCTAATTGTGACGCAGTTAGAGTAATTTTAAATCAGCTTCTTGTCTT	1989
Sbjct	590	CACGCACTAATTAGCTAATTGTGACGCAGTTAGAGTAATTTTAAATCAGCTTCTTGTCTT	649

Query	1990	TTCTTTTCGCAGGTGCTAACTGGGTCTTCCCTGAAGCAGAAGGAACAGTATTTCTACCTT	2049
Sbjct	650	TTCTTTTCGCANGTGCTAACTGGGTCTTCCCTGAAGCAGANNGAACAGTATTTCTACCTT	709
Query	2050	TCCGAACCGAA	2060
Sbjct	710	TCCGAACCGAA	720

I-736/+59Smad1.pGL2

Query	1263	GCTCTGAGAACGGGACGAGTGCCCTCAGCCTTCTGAGGCACTGAGCGGGCTCCTCCACCC	1322
Sbjct	50	GCTCTGAGAACGGGACGAGTGCCCTCAGCCTTCTGAGGCACTGAGCGGGCTCCTCCACCC	109
Query	1323	ACAGGAGCTTATCTTATCCTTCAAATGATTTTCTCTGTACACAAAACCTCTGGCTTTGC	1382
Sbjct	110	ACAGGAGCTTATCTTATCCTTCAAATGATTTTCTCTGTACACAAAACCTCTGGCTTTGC	169
Query	1383	CGATGAATTTGGTCGCATTTTACGCTGTAGTCTTCTGTACCTGTGCGGCTCCTGTGTATC	1442
Sbjct	170	CGATGAATTTGGTCGCATTTTACGCTGTAGTCTTCTGTACCTGTGCGGCTCCTGTGTACC	229
Query	1443	AGTCCATCTCCCTCCAGGAGTGTCGGTTGCTGGAGTGGTTAGCTTCTTTGCAGTTTGT	1502
Sbjct	230	AGTCCATCTCCCTCCAGGAGTGTCGGTTGCTGGAGTGGTTAGCTTCTTTGCAGTTTGT	289
Query	1503	GCTTGGTATCTGTTGGGGCTTGTCACTTTCGTTTACATATGAGGTGGTAATATGCAAATC	1562
Sbjct	290	GCTTGGTATCTGTTGGGGCTTGTCACTTTCGTTTACATATGAGGTGGTAATATGCAAATC	349
Query	1563	AGCCAAGTGGATTCTGTGCCCCGAAACACTCGCATGACCTTGTCCAGGGCTTGGGTGCT	1622
Sbjct	350	AGCCAAGTGGATTCTGTGCCCCGAAACACTCGCATGACCTTGTCCAGGGCTTGGGTGCT	409
Query	1623	GTTGGGAGCAATCTATTCCCTTCTCTGTAAAGCTGCCAATCCTGTTGCCTCAAGCCCC	1682
Sbjct	410	GTTGGGAGCAATCTATTCCCTTCTCTGTAAAGCTGCCAATCCTGTTGCCTCAAGCCCC	469
Query	1683	ACCCACCCTTATAATCTAACTTAACCCTAGTGAGTTACTTAGAGGCTCCAAGCCCAAATA	1742
Sbjct	470	ACCCACCCTTATAATCTAACTTAACCCTAGTGAGTTACTTAGAGGCTCCAAGCCCAAATA	529
Query	1743	CCAGTCGGTCGCAGTGGGTGCTCTGCCTTTAACTGGTGAACCTTAGGAGAACGTAGCTCA	1802
Sbjct	530	CCAGTCGGTCGCAGTGGGTGCTCTGCCTTTAGCTGGTGAACCTTAGGAGAACGTAGCTCA	589
Query	1803	GTTCTGGTGTACAGGACACAAATGTTTCATGAACTCAGCATGGGGTTGCTCTCTAATGA	1862
Sbjct	590	GTTCTGGTGTACAGGACACAAATGTTTCATGAACTCAGCATGGGGTTGCTCTCTAATGA	649
Query	1863	AGTCACCGAATTGGTCTAAGCCAAGCGTGGACCTCTACAGGGTCAGACTTAAGTAGTTAG	1922

Sbjct	650	AGTCACCGAATTGGTCTAAGCCAAGCGTGGACCTCTACAGGGTCAGACTTAAGTAGTTAG	709
Query	1923	TACCTGGCACGCACTAATTAGCTAATTGTGACGCAGTTAGAGTAATTTTAAATCAGCTTC	1982
Sbjct	710	TACCTGGCACGCACTAATTAGCTAATTGTGACGCAGTTAGAGTAATTTTAAATCAGCTTC	769
Query	1983	TTGTCTTTTCTTTTCGCAGGTGCTAACTGGGTCTTCCCTGAAGCAGAAGGAACAGTATTTC	2042
Sbjct	770	TTGTCTTTTCTTTTCGCNNGTGCTAACTGGGTCTTCCCTGAAGCAGAA-GNACAGTATTTC	828
Query	2043	CTACCTTTCCGAACCGAA	2060
Sbjct	829	CTACCTTTCCGAANCGAA	846

I-1109+59Smad1.pGL2

Query	892	TTTCGGGCTGGGGATTAGCTCAGTGGTAGAGCGCTTGCCTAGGAAGCGCAAGGCCCTGG	951
Sbjct	49	TTTCGGGCTGGGGATTAGCTCAGTGGTAGAGCGCTTGCCTAGGAAGCGCAAGGCCCTGG	108
Query	952	GTTTCGGTCCCCAGTTCGAAAAAAGAACTAAAAAAAAAAAAAAAAAATTATGGTTGTT	1011
Sbjct	109	GTTTCGGTCCCCAGTTCGAAAAAAGAACTAAAAA---AAAAAAAAAATTATGGTTGTT	165
Query	1012	TTCAACCACCATATTATTGTTGATGTCTTAGTTATACTAAGGCCGAGGGGTCTTGGTTAT	1071
Sbjct	166	TTCAACCACCATATTATTGTTGATGTCTTAGTTATACTAAGGCCGAGGGGTCTTGGTTAT	225
Query	1072	TACCATTATTATACGTTTACACACAGCTACGTGCACAGACGTCGAGTTGTCTGCCCTGG	1131
Sbjct	226	TACCATTATTATACGTTTACACACAGCTACGTGCACAGACGTCGAGTTGTCTGCCCTGG	285
Query	1132	TTGTTTGTTCATTATGGAGACAGGGTCTCTGGCTGACCTGAGCTTGCTAATGAACTGGTC	1191
Sbjct	286	TTGTTTGTTCATTATGGAGACAGGGTCTCTGGCTGACCTGAGCTTGCTAATGAACTGGTC	345
Query	1192	CAGCCAGCAAGTTCAGAGATCCGCCTGTGCCTTCTCAGAGTGGGCAGACAAGCACGGTCC	1251
Sbjct	346	CAGCCAGCAAGTTCAGAGATCCGCCTGTGCCTTCTCAGAGTGGGCAGACAAGCACGGTCC	405
Query	1252	GCTCAGTGTGGGCTCTGAGAACGGGACGAGTGCCCTCAGCCTTCTGAGGCACTGAGCGGG	1311
Sbjct	406	GCTCAGTGTGAGCTCTGAGAACGGGACGAGTGCCCTCAGCCTTCTGAGGCACTGAGCGGG	465
Query	1312	CTCCTCCACCCACAGGAGCTTATCTTATCCTTCAAATGATTTTCTCTGTACCAAAACC	1371
Sbjct	466	CTCCTCCACCCACAGGAGCTTATCTTATCCTTCAAATGATTTTCTCTGTACCAAAACC	525
Query	1372	TCTGGCTTTGCCGATGAATTTGGTCGCATTTTACGCTGTAGTCTTCTGTACCTGTGCGG	1431
Sbjct	526	TCTGGCTTTGCCGATGAATTTGGTCGCATTTTACGCTGTAGTCTTCTGTACCTGTGCGG	585

Query	1432	TCCTGTGTATCAGTCCATCTCCCTCCAGGAGTGTGGTGTGCTGGAGTGGTTAGCTTCTT	1491
Sbjct	586	TCCTGTGTATCAGTCCATCTCCCTCCAGGAGTGTGGTGTGCTGGAGTGGTTAGCTTCTT	645
Query	1492	TGCAGTTTTGTGCTTGGTATCTGTTGGGGCTTGTCACTTTCGTTTACATATGAGGTGGTA	1551
Sbjct	646	TGCAGTTTTGTGCTTGGTATCTGTTGGGGCTTGTCACTTTCGTTTACATATGAGGTGGTA	705
Query	1552	ATATGCAAATCAGCCAAGTGGATTCTGTGCCCCGAAACACTCGCATGACCTTTGTCCAGG	1611
Sbjct	706	ATATGCAAATCAGCCAAGTGGATTCTGTGCCCCGAAACACTCGCATGACCTTTGTCCAGG	765
Query	1612	GCTTGGGTGCTGTTGGGAGCAATCTATTCCCTTCCTCTTGTAAGCTGCCAATCCTGTTG	1671
Sbjct	766	GCTTGGGTGCTGTTGGGAGCAATCTATTCCCTTCCTCTTGTAAGCTGCCAATCCTGTTG	825
Query	1672	CCTCAAGCCCCACCCACCCTTATAATCTAACTTAACCCTAGTGAGTTACTTAGAGGCTCC	1731
Sbjct	826	CCTCAAGCCCCACCCACCCTTATAATCTAACTTAACCCTAGTGAGTTACTTAGAGGCTCC	885
Query	1732	AAGCCCAAATACCAGTCGGTCGCAGTGGGTGCTCTGCCTTTAACTGGTGAACCTTAGGAG	1791
Sbjct	886	AAGCCCAAATACCAGTCGGTCACAGT-GGNGCTCTGCCTTTAACTGGTGAACCTT-NNAN	943
Query	1792	AACGTAGCTCAGTTCCTGGTGTACAGGACACAAATGT	1829
Sbjct	944	NACGTAGCTCAGTTCCT-GNGTCACAGGACACAAATGT	980
Gap			
Query	1967	ATTTTAAATCAGCTTCTTGTCTTTCTTTTCGCAGGTGCTA	2006
Sbjct	1094	ATTTTAANNCAGCTTCTTGTCTTTCTTTNGCAGGTGCTA	1133

I-2015+59Smad1.pGL2 (partial)

Query	1514	GTTGGGGCTTGTCACCTTCGTTTACATATGAGGTGGTAATATGCAAATCAGCCAAGTGGTA	1573
Sbjct	55	GTTGGGGCTTGTCACCTTCGTTNNNTATGAGGTGGGAATATGCAAATCAGCCAAGTGGTA	114
Query	1574	TTTCTGTGCCCCGAAACACTCGCATGACCTTTGTCCAGGGCTTGGGTGCTGTTGGGAGCAA	1633
Sbjct	115	TTTCTGTGCCCCATACACTCGCATGACCTTTGTCCAAGGCTTGGGTGCTGTTGGGAGCCA	174
Query	1634	TCTATTCCCTTCCTCTTGTAAGCTGCCAATCCTGTTGCCTCAAGCCCCACCCACCCTTA	1693
Sbjct	175	TCTATTCCCTTCCTCTGTNNAGCTGCCAATCCTGTTGCCTCAAGCCCCACCCACCCTTA	234
Query	1694	TAATCTAACTTAACCCTAGTGAGTTACTTAGAGGCTCCAAGCCCAAATACCAGTCGGTCG	1753
Sbjct	235	TAATCTAACTTACCCCTAGTGAGTTACTTAGAGGCTCCAAGCCCATACCGGACGGCCC	294

Query	1754	CAGTGGGTGCTCTGCCTTTAACTGGTGAACCTTTAGGAGAACGTAGCTCAGTTCCTGGTGT	1813
Sbjct	295	CACCGGGTCTCTGCCTTTAACTGGTGAATTTTAGGAGAACGTNGCTCNGTTCCTGGTG-	353
Query	1814	CACAGGACACAAATGTTTCATGAACCTCAGCATGGGGTTGCTCTCTAATGAAGTCACCGAAT	1873
Sbjct	354	GACAGGACACAAATGTTTCATGAAGTCACCATGGGGTTGCTCTCTAATGAAGTCCCCAAAT	413
Query	1874	TGGTCTAAGCCAAGCGTGGACCTCTACAGGGTCAGACTTAAGTAGTTAGTACCTGGCACG	1933
Sbjct	414	TGGTCTACNCCAAGCGTGGACCTCCCCCGGCCAGACTTATTTGTTAGTACCTGGTGCG	473
Query	1934	CACTAATTAGCTAATTGTGACGCAGTTAGAGTAATTTTAAATCAGCTTCTTGTCTTTTCT	1993
Sbjct	474	CTCTAATTACCTAATTGAGACGCAGGTACATNGATTTTAAATCACCTTCTTGTCTTTTCT	533
Query	1994	TTCGCAGGTGCTAACTGGGTCTTCCCTGAAGCAGAAGGAACAGTATTTCCTACCTTTCCG	2053
Sbjct	534	TTCGCANGTGCTAACTGGGTCTTCCCTGAATGAGAAGGTACCGTATTTCCTACCNITCCT	593
Query	2054	AACCGA	2059
Sbjct	594	AANCGA	599

Fig.23. Sequences alignment. Sequences were studied using online tool from NCBI (<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi>). Rat genomic DNA sequences containing Smad1 locus F-1200/+157 and I-2000/+180 were obtained from Ensembl (<http://www.ensembl.org/Rattusnorvegicus/geneview?db=core;gene=ENSRNOG00000018483>).

In the analyses, genomic templates were registered as 'Query'; testing sequences were registered as 'Sbjct'. Sequences from most of the constructs were completely matched to the templates while a few with base pair changes including F-605/+93Smad1.pGL2, I-610/+59Smad1.pGL2, I-736/+59Smad1.pGL2, I-1109+59Smad1.pGL2. The changes could result from individual diversity, alteration from cloning process or inaccuracy of DNA sequencing. Construct I-2015+59Smad1.pGL2, was partially sequenced and the quality of the sequencing was not good, which could result from its long length.

3. Transcriptional activity Analyses of rat Smad1 gene

To analyze the transcriptional activity of the cloned 5'-flanking region and intron1 region, four sets of experiments were performed. Fragments from rat Smad1 5'-flanking region and intron1 region were fused to firefly luciferase gene to generate Smad1.pGL2 constructs (Fig.15) which subsequently transfected into Chang liver cells or CFSC-8B cells to analyze the transcriptional activity. Each transfection was carried out in triplicate and repeated at least twice. Data were presented as means $\pm$ S.D. from at least six determinations.

The promoterless pGL2-Basic vector (-pGL2) was used to transfect cells as a negative control. A hybrid luciferase gene directed by mouse typeX promoter (CollagenX.pGL2, Fig.24), which kindly provided by Dr. Hong Shen (University of Manitoba, Canada), was used as a positive control. For DNA uptake, pRL-TK vector (Fig.25), which expresses Renilla luciferase, was co-transfected into cells and used subsequently to correct transfection efficiency. Therefore, transcriptional activity was measured by firefly luciferase activity over Renilla luciferase activity.

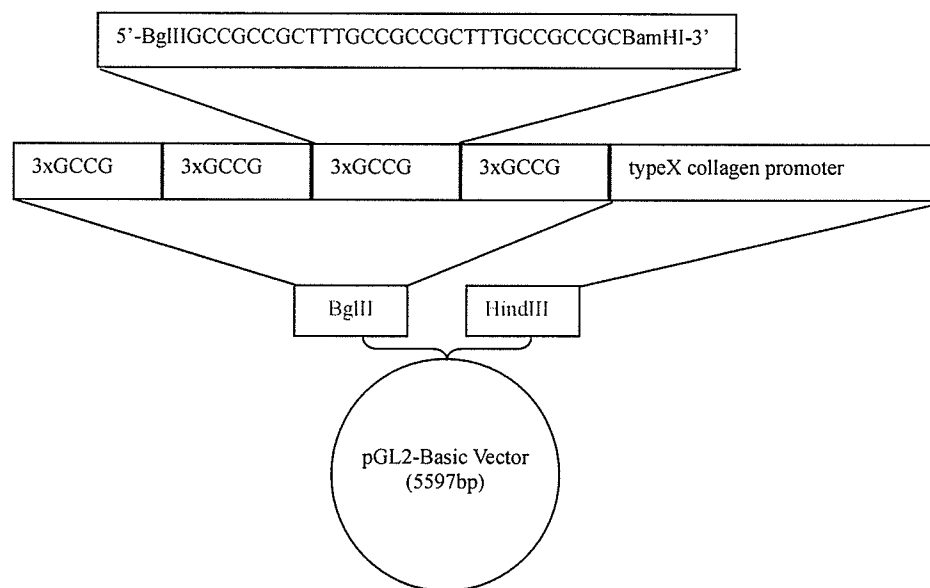


Fig.24. CollagenX.pGL2 map.

**This item has been removed
due to copyright issues. To
view it, refer to its source.**

Fig.25. pRL-TK vector map.

3.1 Identification of rat *Smad1* promoter regions

To investigate the transcriptional activity of rat *Smad1* gene, constructs F-1005/+93*Smad1*.pGL2 and I-2015/+59*Smad1*.pGL2 with the longest fragments from rat *Smad1* 5'-flanking region and intron1 region respectively were generated. Their activities to promote expression of the reporter gene were evaluated in Chang liver cells. -pGL2 was used as negative control and CollagenX.pGL2 was used as positive control.

Transcriptional activity was identified in 5'-flanking region. As shown in Figure26-A, a marked increase in activity was observed in the cells transfected with F-1005/+93*Smad1*.pGL2 in comparison with negative control. The activity of F-1005/+93*Smad1*.pGL2 was approximately 2-fold greater than positive control. Transcriptional activity was identified in intron1 region as well. In Figure26-B, significant activity was observed with I-2015/+59*Smad1*.pGL2 compared with the negative control and the activity of I-2015/+59*Smad1*.pGL2 was 5-fold lower than positive control.

The identification of transcriptional activities in rat *Smad1* 5'-flanking region and intron1 region indicated that rat *Smad1* had two promoter regions. Both of them could initiate *Smad1* transcription. Additionally, we found that the activity of F-1005/+93*Smad1*.pGL2 was 10-fold higher than that of I-2015/+59*Smad1*.pGL2, which indicated that 5'-flanking region contained predominant promoter. Taking into account the fact that the exon1 of rat *Smad1* gene is noncoding, the second promoter appeared in intron1 region might be the backup mechanism for this important gene being expressed.

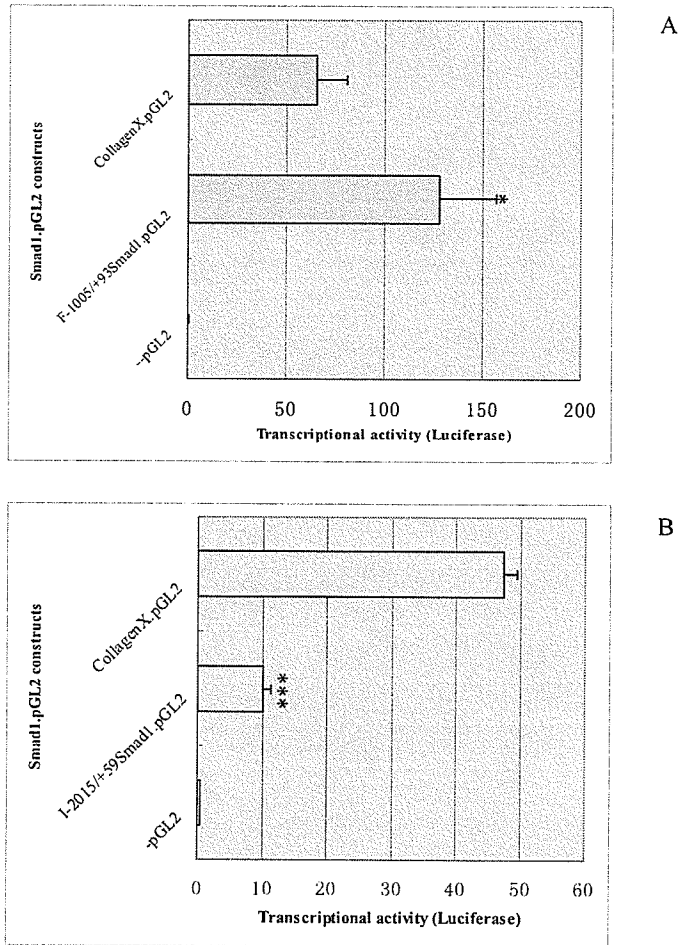


Fig.26. Transcriptional activity analyses of rat Smad1 5'-flanking region and intron1 region in Chang liver cells. F-1005/+93Smad1.pGL2 and I-2015/+59Smad1.pGL2 were transfected into Chang liver cells. A: Transcriptional activity analysis of 5'-flanking region. Great activity was observed with F-1005/+93Smad1.pGL2 compared with negative control (-pGL2) and the activity of this construct was 2-fold higher than that of the positive control (CollagenX.pGL2). B: Transcriptional activity analysis of intron1 region. Activity of I-2015/+59Smad1.pGL2 was observed compared with negative control and 5-fold lower than that of positive control. * indicated $P < 0.05$ and *** indicated $P < 0.001$ between compared groups.

3.2 Deletional analyses of rat Smad1 5'-flanking region and intron1 region

To define the elements involved in regulation of rat Smad1 gene expression, twelve 5'-end deletional fragments from both identified promoter regions were introduced upstream of the translation start site of the luciferase reporter sequence in pGL2-Basic vector to generate a series of rat Smad1.pGL2 constructs. These constructs were transiently transfected into Chang liver cells and their activities were investigated. All the constructs demonstrated significant promoter activities compared with the negative control -pGL2.

As shown in Figure27-A, F-1005/+93Smad1.pGL2 showed the highest transcriptional activity among the five constructs carrying deletional fragments from Smad1 5'-flanking region. Deletions from F-1005bp to F-163bp had little effect on the activity. However, a further deletion to F-56bp resulted in a 4.5-fold decrease on activity, which suggested the presence of a positive regulatory element located between F-56 and F-163 in the 5'-flanking region of rat Smad1 gene.

In Figure27-B, I-2015/+59Smad1.pGL2 showed lower transcriptional activity compared with the other six constructs containing deletional fragments from Smad1 intron1 region. Deletion from I-2015bp to I-1109bp resulted in a 4.7-fold increase on the activity. An additional deletion to I-389bp did not show much difference. However, the removal of another 117bp sequence resulted in a big loss of the activity. These results suggested that the presence of two different regulatory elements: a negative regulatory element located between I-1109bp and I-2015bp; and a positive regulatory element located between I-273 and I-389.

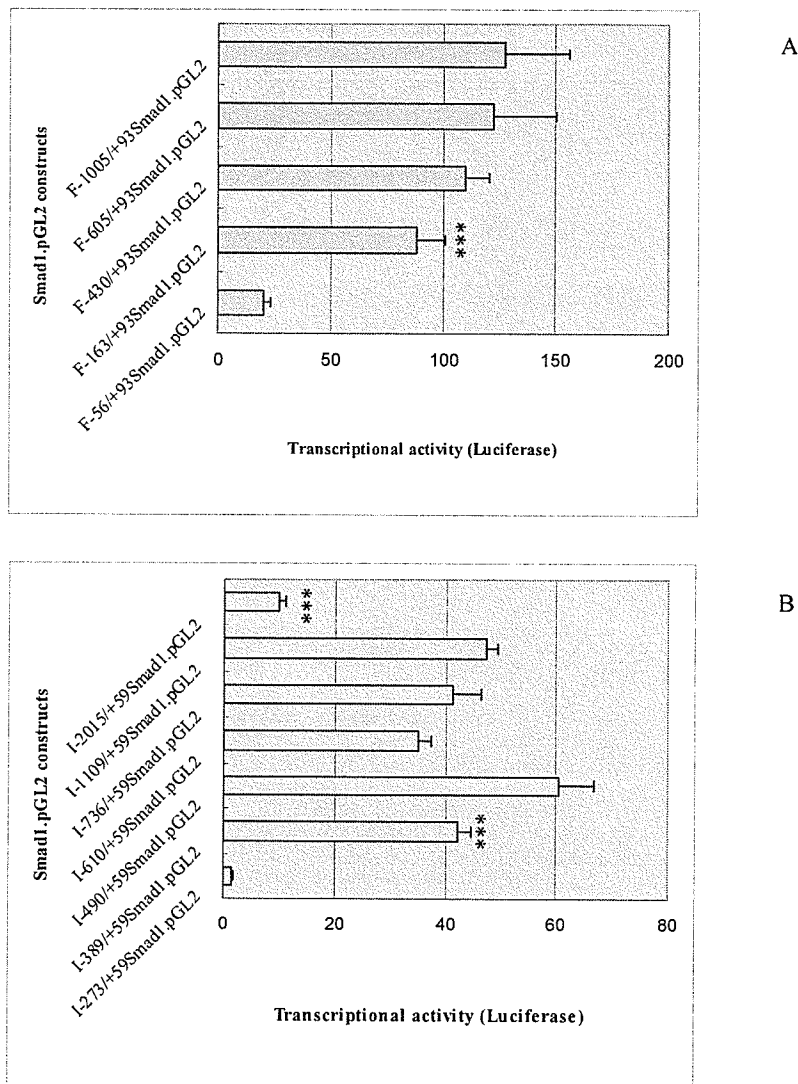


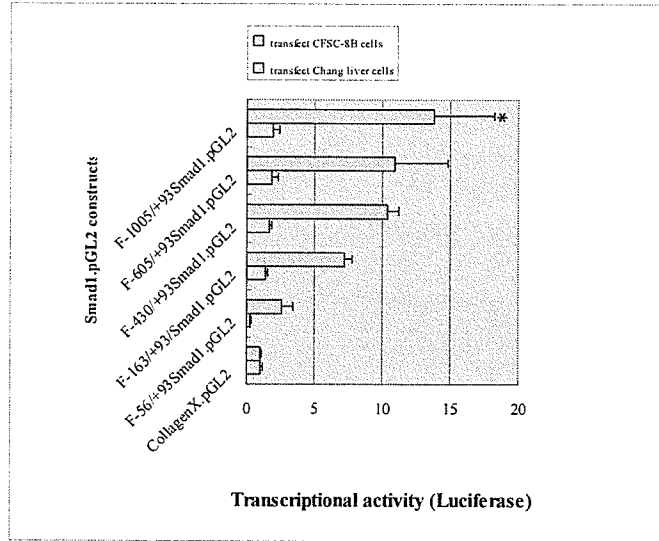
Fig.27. Deletional analyses of rat Smad1 5'-flanking region and intron1 region. The activities of the constructs were shown as mean $\pm$ S.D after subtraction of -pGL2 activity. A: Deletion of 5'-flanking region. Deletions from F-1005 to F-163 had little effect on the activity but an additional deletion to F-56 resulted in a 4.5-fold decrease. B: Deletion of intron1 region. Deletion from I-2015bp to I-1109bp resulted in a 4.7-fold increase on the activity. An additional deletion from I-389bp to I-273bp resulted in a big loss of the activity. *** indicated $P < 0.001$ between the compared groups.

3.3 Comparison of the activities in Chang liver cells and CFSC-8B cells

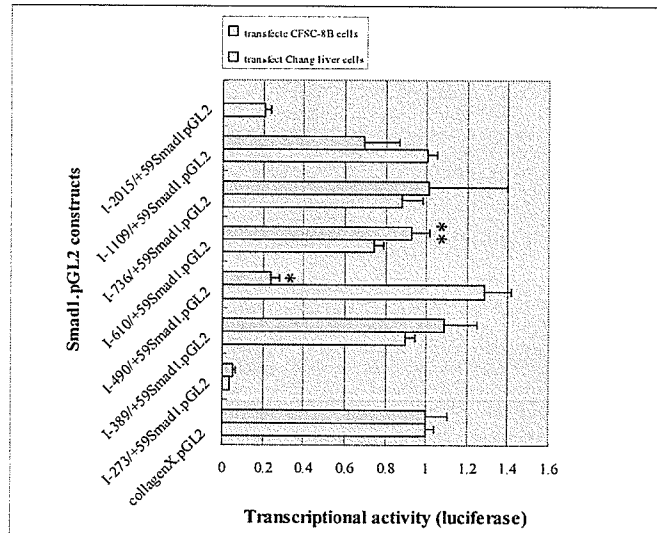
To investigate transcriptional activities of 5'-flanking and intron1 regions of rat Smad1 gene in activated hepatic stellate cells, twelve deletional constructs and CollagenX.pGL2 were transiently transfected into CFSC-8B cells (Cirrhotic Fat-Storing Cells-8B). Transcriptional activities in CFSC-8B cells were compared with the activities in Chang liver cells. For comparison, the activity of CollagenX.pGL2 was set up as 1 and the activities of all these testing constructs were expressed as a percentage of CollagenX.pGL2.

As shown in Figure28-A, the constructs with fragments from 5'-flanking region showed same pattern in both cell lines. However, generally, the activities of these constructs in CFSC-8B cells were much higher than in Chang liver cells. The activity of F-1005/+93Smad1.pGL2, which contained all the possible regulatory elements in testing 5'-flanking region, was 7-fold greater in CFSC-8B cells than in Chang liver cells.

For intron1 region (Fig.28-B), six constructs were studied in both cell lines, but I-2015/+56Smad1.pGL2 was only checked in Chang liver cells. The general activities of these constructs in Chang liver cells and CFSC-8B cells showed little difference. In Chang liver cells, the activities of I-610/+53Smad1.pGL2, I-490/+53Smad1.pGL2 and I-389/+53Smad1.pGL2 were at same level, However, in CFSC-8B cells, the activity of I-610/+53Smad1.pGL2 was 4-fold greater than that of I-490/+53Smad1.pGL2 and the activity of I-490/+53Smad1.pGL2 was 6-fold lower than that of I-389/+53Smad1.pGL2. This result suggested that specific transcriptional factors from CFSC-8B cells maybe contribute to the binding of sequences I-610/-490 and I-490/-389 and involve in regulation of gene expression.



A



B

Fig.28. Comparison of activities in Chang liver cells and CFSC-8B cells. The activity of CollagenX.pGL2 was set up as 1 and the activities of all the testing constructs were expressed as a percentage of CollagenX.pGL2. A: Comparison of 5'-flanking region. The general activities of the constructs in CFSC-8B cells were much higher than in Chang liver cells. B: Comparison of intron1 region. The general activities of these constructs in both cell types showed little difference. However, sequence from I-389bp to I-610bp got different response in different cell types. * indicated $P < 0.05$ and ** indicated $P < 0.01$ between the compared groups.

3.4 Comparison of the activities in activated and quiescent CFSC-8B cells

3.4.1 Morphology and viability of CFSC-8B cells cultured on MatrigelTM

CFSC-8B cell line is an activated rat hepatic stellate cell line. When CFSC-8B cells were cultured on uncoated plastic dishes, cells had the characteristic enlarged myofibroblastic appearance (Fig.29-A, B). To obtain quiescent hepatic stellate cells, CFSC-8B cells were cultured on BD MatrigelTM Basement Membrane Matrix. MatrigelTM is a solubilized basement membrane preparation extracted from the Engelbreth-Holm-Swarm (EHS) mouse sarcoma. The major components of MatrigelTM are laminin and collagene IV. CFSC-8B cells were seeded on MatrigelTM coated dishes and cultured for 1 day or 3 days. As shown in Fig.26-C, D, cells plated on MatrigelTM substrate formed clusters connected by a filamentous network and the proliferation of these cells remained very slow.

CFSC-8B cells were cultured on plastic and MatrigelTM substrate respectively and were harvested on day1 and day3. Each experiment was performed in duplicates. The expression of α -Smooth muscle actin (α -SMA), the marker of activated hepatic stellate cells, was detected by Western Blotting and the results were analyzed by Densitometry Study. As shown in Fig.30, expression of α -SMA in CFSC-8B cells cultured on plastic was intense. However, α -SMA expression was downregulated by culturing cells on MatrigelTM substrate on a time-dependent manner. On day1, α -SMA expression was suppressed by 70% and on day3 the expression was suppressed by 90%.

The change of cell morphology and the loss of α -SMA expression demonstrated that the activated hepatic stellate cells, CFSC-8B cells, were reverted to quiescent stage by culturing on MatrigelTM substrate.

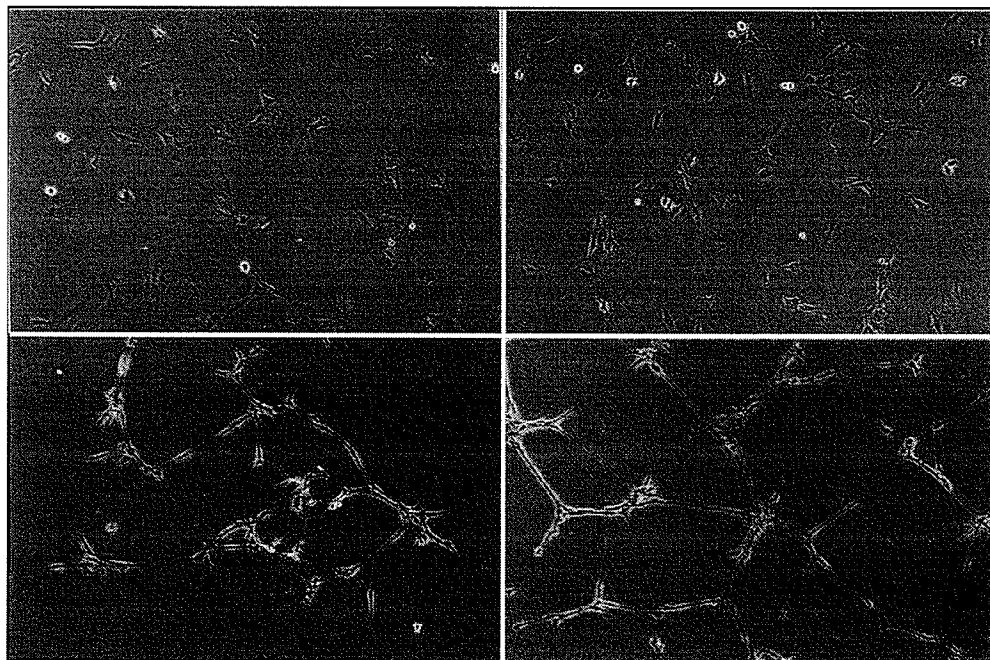


Fig.29. Morphology of CFSC-8B cells cultured on plastic dish and MatrigelTM dish. A and B: Cells cultured on plastic dish for 1 day and 3 days. Cells had the characteristic myofibroblastic appearance. C and D: Cells cultured on MatrigelTM for 1 day and 3 days. Cells formed clusters connected by a filamentous network and the proliferation of these cells remained very slow.

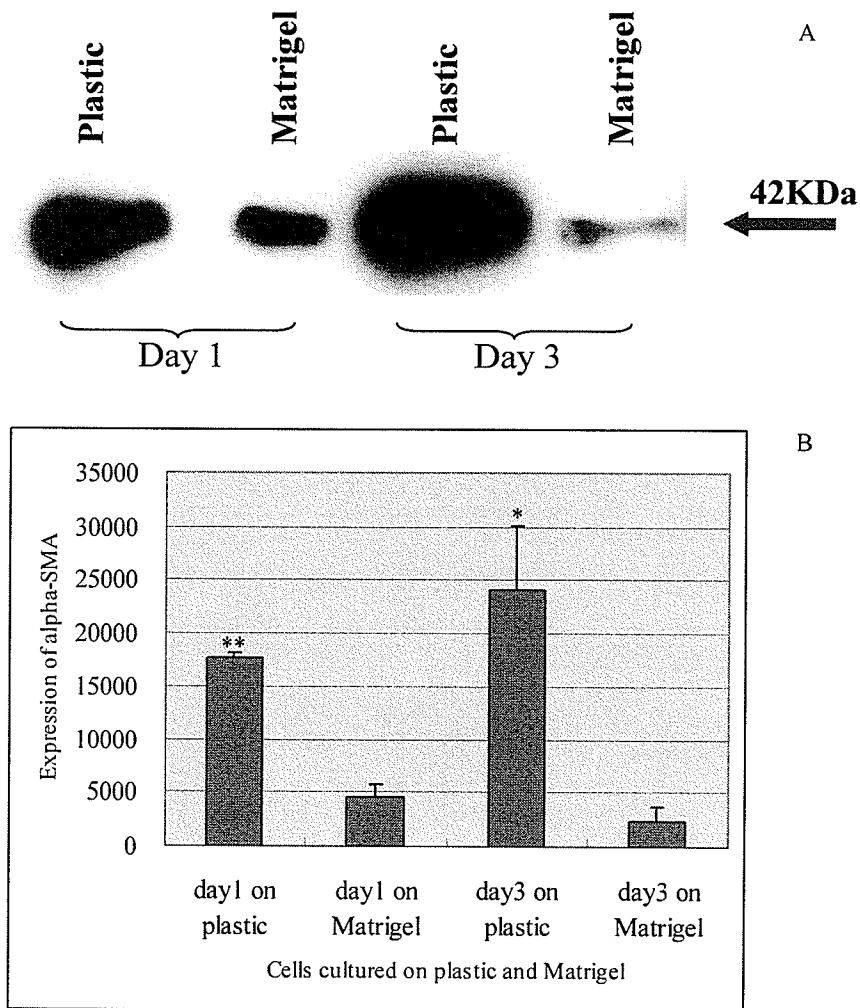


Fig.30. Expression of α -Smooth muscle actin (α -SMA). A: Western Blotting analysis of α -SMA expression. CFSC-8B cells cultured on plastic and MatrigelTM substrate were harvested on day1 and day3. α -SMA expression increased on a time-dependent manner when cells cultured on plastic but was downregulated by culturing on MatrigelTM substrate. B: Densitometry analysis of α -SMA expression. On day1, α -SMA expression was suppressed by 70% and on day3 the expression was suppressed by 90%. Data were presented as means $\pm$ S.D. from two determinations. * indicated $P < 0.05$ and ** indicated $P < 0.01$ between the compared groups.

3.4.2 Transcriptional activity of rat *Smad1* in quiescent CFSC-8B cells

To analyze the transcriptional activities of two promoter regions of rat *Smad1* gene in quiescent hepatic stellate cells, five constructs of 5'-flanking region and five constructs of intron1 region were transfected into CFSC-8B cells cultured on MatrigelTM substrate on day3. Each transfection reaction was carried out in triplicates. -pGL2 was used as a negative control. As shown in Fig.31, the activities of those constructs in quiescent CFSC-8B cells were lower than the activity of negative control which indicated that in quiescent CFSC-8B cells, the function of two promoter regions of rat *Smad1* gene was repressed.

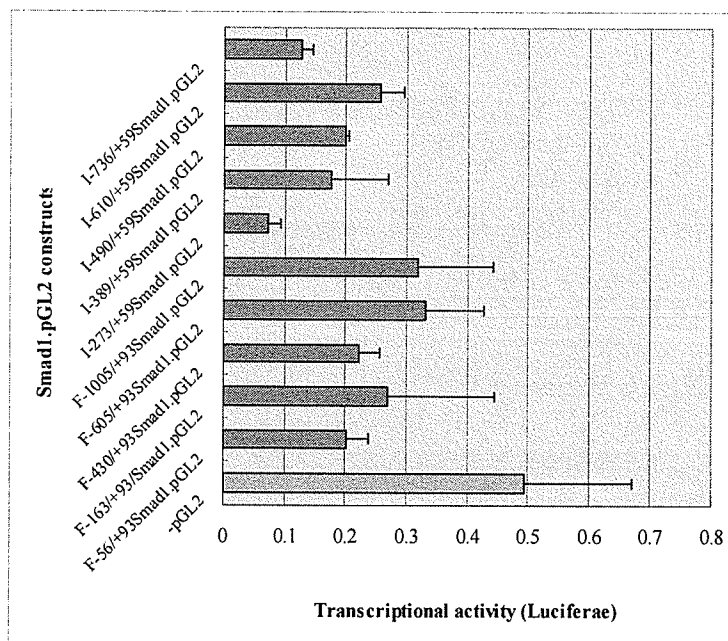


Fig.31. Transcriptional activities of rat *Smad1* in quiescent CFSC-8B cells. -pGL2 was used as a negative control and data were presented as means $\pm$ S.D. from three determinations. The activities of those constructs in quiescent CFSC-8B cells were lower than that of negative control.

4. Investigation of transcriptional factors in rat Smad1 5'-flanking region

4.1 Characterization of rat Smad1 5'-flanking region

In order to investigate the 5'-flanking region of rat Smad1 gene, a sequence of F-430/+1 was analyzed in search of the binding sites for possible transcriptional factors by using the online tool in TRANSFAC database (Yutaka Akiyama, 1995). Searching was restricted in vertebrate and the results were selected by scoring. Results with a higher score than 85 were displayed in Fig.32. GC-boxes, which were the result with a score of 83, were also shown in this figure.

This region did not contain either TATA-box or Inr. However, sequence analysis revealed that F-163/+1 sequence was a CG-rich sequence. CG content was up to 70%. Two GC-boxes, the binding sites for Sp1, were identified in the sequence of F-163/-56 and F-56/+1 respectively. A C/EBP-like motif, the binding site for CCAAT/enhancer binding protein (C/EBP), was found in F-163/-56 as well.

These findings might indicate that rat Smad1 5'-flanking region contained a CG-rich promoter, which could contribute to the transcriptional activity and the second GC-box and the C/EBP motif might be the positive regulatory elements identified between F-163 and F-56 in transcriptional activity analysis.

Several other transcriptional factor binding sites were identified in F-430/+1 sequence, including the binding sites for CdxA, CREB and GATA factors. These potential transcriptional factors and their types were listed in Table5.

Database: TRANSFAC MATRIX TABLE, Rel.3.3 06-01-1998
 Query: F-430/+1 (430 bases)
 Taxonomy: Vertebrate
 Threshold: 85.0 point

TFMATRIX entries with High-scoring:

```

1 GCTTTTGTCA TTGCCGCAG CAAGGGTACA TACTAGAAAT AAATGCATT entry      score
  ↳ F-430 ----->          <----- M00101 CdxA  94.3
                                <----- M00050 E2F   90.8
                                <----- M00101 CdxA  90.0
                                <----- M00100 CdxA  89.7
                                -----> M00101 CdxA  86.4
                                <----- M00100 CdxA  85.9

51 ACCTAGCAGT TAACAGGTCA CCCAGTTTGG GGCTCAGGCC ACATGGAGGG entry      score
                                -----> M00217 USF   86.4
                                -----> M00050 E2F   86.2

101 CCGGTAAATC ACTGGGTGAG TGATTGCAGG GCCATTAAAA TGGGTTTTTA entry      score
                                <----- M00101 CdxA  92.1
                                <----- M00173 AP-1  89.7

151 GGACGATTGG GTGGGTTTGA TGGTCGTGTT TGAGAACATT TCAGAAAATA entry      score
  ----->                                M00075 GATA-1 91.8
                                <----- M00109 C/EBPb 88.7
  ----->                                M00077 GATA-3 87.5
                                <----- M00101 CdxA  86.4
                                ----- M00101 CdxA  86.4
  ----->                                M00076 GATA-2 86.2

201 ACACAGGCTT TGGAACTTGA CTGGCAGCTA GGCAGGGACT ACAGAGTCAG entry      score
  -->                                M00101 CdxA  86.4
                                <-- F-163 M00240 Nkx-2. 86.0
                                ↳ C/EBP motif

251 AGGCAGGCGC GGGCCGGGAA TTTGGAGAGG ATCCCTAAGC AGCAGCGGGC entry      score
                                -----> M00159 C/EBP  86.9
                                GC-box2

301 GCGCGCGGTG CGCGTGTGAG CGAGGGCGAG GGGCGGGCGG GTTGTGCGCC entry      score
                                ↳ F-56
                                GC-box1

351 GGGTATTGGC AGGCGGGCTG TGGAGGCTGG GCGGTTCCGA CTGCCTGACG entry      score
                                ----- M00039 CREB  87.1
                                ----- M00041 CRE-BP 86.2
                                <----- M00041 CRE-BP 86.2
                                <----- M00039 CREB  85.6
                                F+1

401 CCAGCGCGAC CAGATCAATC CAGGCTCGGG entry      score
  <----- M00106 CDP CR 95.0
  --> M00039 CREB  87.1
  -----> M00096 Pbx-1 86.3
  --> M00041 CRE-BP 86.2
  --- M00041 CRE-BP 86.2
  --- M00039 CREB  85.6
  
```

Total 24 high-scoring sites found.
 Max score: 95.0 point, Min score: 85.6 point

Fig.32. Sequence analysis of F-430/+1 of rat Smad1 5'-flanking region.

Table5. Potential transcriptional factors for F-430/+1 sequence

Sequence	Potential transcription factors	
	<i>Factors</i>	<i>Type</i>
F-430/-163	CdxA	Caudal-type homeodomain protein
	E2F	control of cell cycle progression and the action of tumor suppressor protein
	USF	upstream stimulatory factor
	AP-1	activated protein-1
	GATAs	globin transcription factors
	C/EBPb	CCAAT/enhancer binding protein-beta
	Nkx-2.	Cardiac-specific homeodomain protein
F-163/-56	C/EBP	CCAAT/enhancer binding protein
	GC-box	Sp1
F-163/+1	CREB	CRE-binding protein
	CREB-BP	cAMP response element binding protein
	CDP DR	cut-like homeodomain protein
	Pbx-1	Co-factor for many homeodomain proteins
	GC-box	Sp1

4.2 Interactions between *trans*-factors and *cis*-elements in 5'-flanking region

To examine the interactions between *trans*-factors and *cis*-elements in rat Smad1 5'-flanking region, Electrophoresis Mobility Shift Assay (EMSA) was performed. Two probes covering different regions (Fig.33) were used in this experiment. Probe1 covering the sequence from F-56 to F+1 of the 5'flanking region, was produced commercially (Invitrogen). Probe2 coding from F-163 to F-56 was generated using PCR (Fig.34) and then gel-purified. Probe information and PCR conditions were shown in Table6 and Table7 respectively. Nuclear proteins were extracted from Chang liver cells (human transformed liver cells), CFSC-8B cells (rat cirrhotic fat-storing cells) and H-4-II-E cells (rat hepatoma cells). In each reaction, 1ng of radiolabeled double-stranded probes was incubated with 5µg of nuclear proteins. Labeled probe without protein interaction was used as negative control. Results were shown in Fig.35.

For Probe1 (F-56/+1, Fig.35-A), interaction with CFSC-8B extracts and with H-4-II-E extracts displayed similar mobility shift pattern, showing formation of similar-sized DNA-protein complex, which indicated the same protein bound to F-56/+1 sequence in CFSC-8B cells and H-4-II-E cells. Interaction with Chang liver extracts displayed two shift bands. Other than the same band shown in Probe1/CFSC-8B extracts interaction, a smaller DNA-protein complex was found, which indicated more proteins involved in the formation of DNA-protein complexes in Chang liver cells. The different binding profiles in human cells and rat cells possibly resulted from species differences.

No clear binding pattern was found in the interactions of Probe2 (F-163/-56, Fig35-B) and nuclear extracts. Experiment was repeated more than twice (with different amount of probes and proteins) and same result was shown.

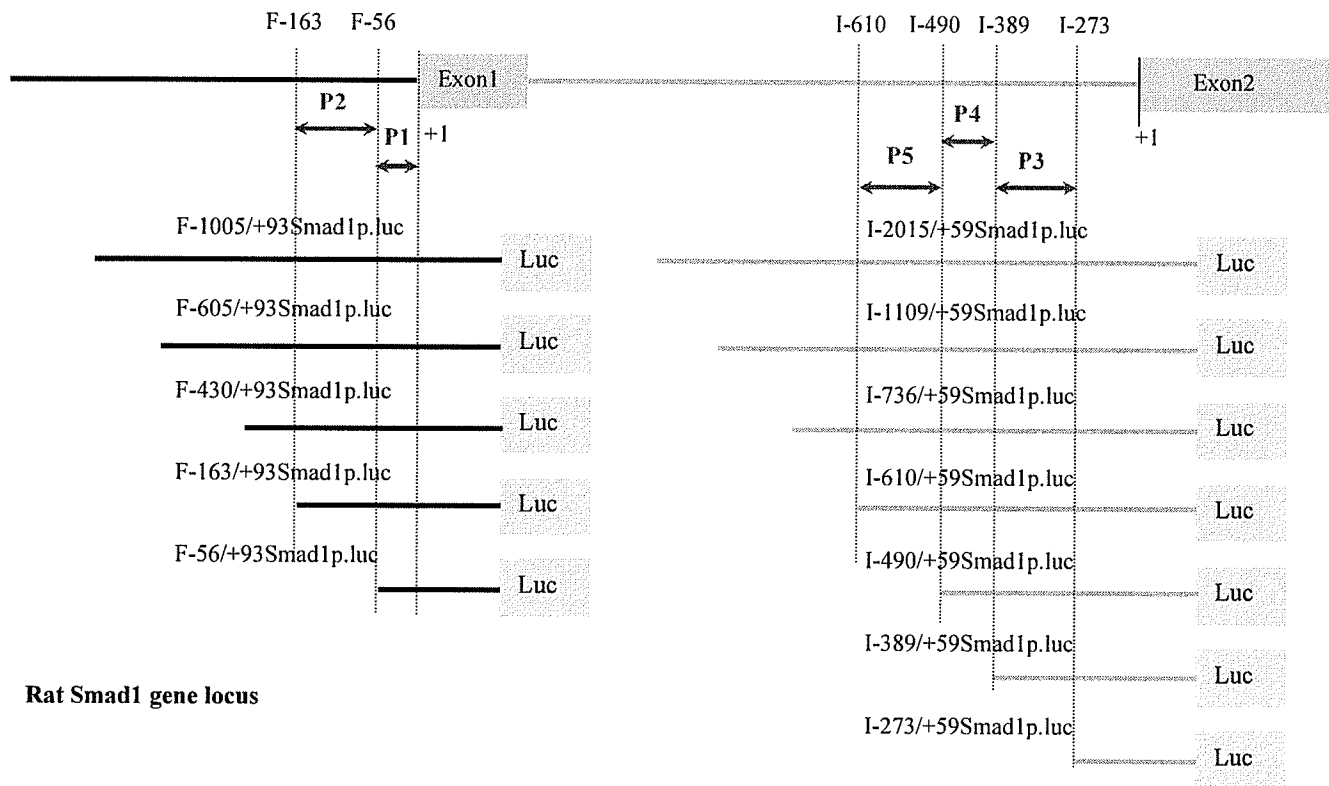


Fig.33. Schematic organization of rat Smad1 probes for EMSA. Pink arrows indicated the position of probes on rat Smad1 gene. The position of probes was also compared with that of Smad1.pGL2 constructs.

Table6. Information of EMSA probes

Name	Sequences	Location	Length
P1	5'-ggctgggcgggtccgactgcctgacgccagcgcgaccaga tcaatccaggctcggg-3'	F-56/+1	56bp
P2	5'-gaatttgagaggatccctaagcagcagcgggcgcgcgc ggtgcgcgtgtgagcgagggcgagcggcggcggggtgt gcgccgggtattggcaggcgggctgtgga-3'	F-163/-56	107bp
P3	5'-gcttgggtgctgttgggagcaatctattcccttctcttgtaaa gctgccaatctgttgctcaagccccaccacccttataatc taacttaaccctagttagttacttagagg-3'	I-389/-273	116bp
P4	5'-tctgttgggctgtcactttcgtttacatatgaggtggaatat gcaaatcagccaagtggatttctgtgcccgaacactcgcat gacctttgtccagg-3'	I-490/-389	101bp
P5	5'-tttggtcgcattttacgctgtagtcttctgtacctgtcggctcct gtgtatcagtcacatctccctccaggagtgtcggttgcaggag ggtttagcttcttgcagttttgtgcttgga-3'	I-610/-490	121bp

Table7. PCR conditions of EMSA probes

Name	Sequences	Temp.	Product
P2-Upper	5'-GAA TTT GGA GAG GA-3'	51°C	107bp
P2-Lower	5'-TCC ACA GCC CGC CT-3'		
P3-Upper	5'-GCT TGG GTG CTG TTG GGA G-3'	58°C	116bp
P3-Lower	5'-CCT CTA AGT AAC TCA CTA GG-3'		
P4-Upper	5'-TCT GTT GGG GCT TGT CAC TT -3'	56°C	101bp
P4-Lower	5'-CCT GGA CAA AGG TCA TGC GA-3'		
P5-Upper	5'-TTT GGT CGC ATT TTT ACG CT-3'	58°C	121bp
P5-Lower	5'-TAC CAA GCA CAA AAC TGC AA-3'		

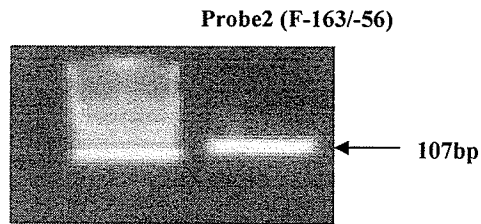
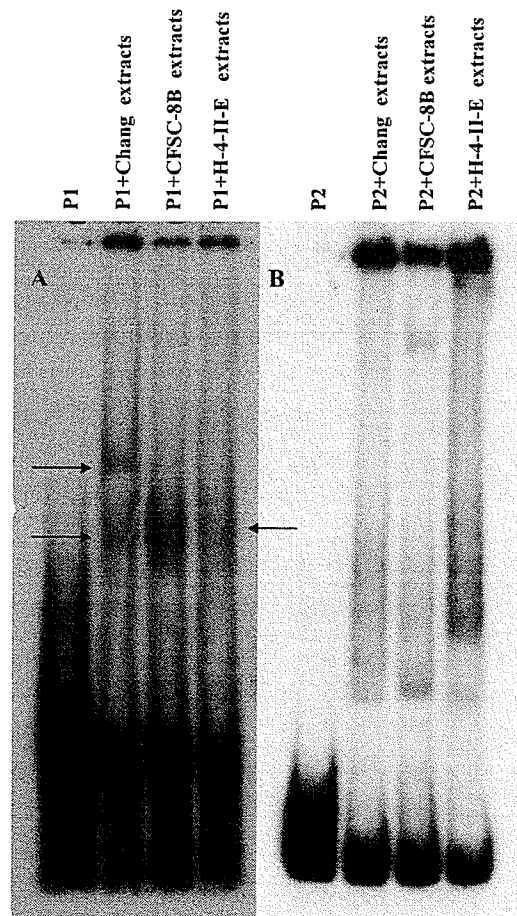


Fig.34. PCR cloning of EMSA probe2. Sequence from F-163 to F-56 of rat Smad1 5'-flanking region was cloned by using primer pair 5'-GAA TTT GGA GAG GA-3' and 5'-TCC ACA GCC CGC CT-3'. PCR was performed at 57°C and the product was gel-purified.

Fig.35. The interaction between nuclear proteins and probes of rat Smad1 5'-flanking region. Nuclear proteins were isolated from Chang liver cells, CFSC-8B cells and H-4-II-E cells. A: Probe1 (F-56/+1) and nuclear extracts interaction. Two DNA-protein binding complexes were detected in interaction of Probe1 and Chang liver extracts and one was shown in interaction with CFSC-8B and H-4-II-E extracts. B: Probe2 (F-163/F-56) and nuclear extracts interaction. No clear pattern was found in Probe2/nuclear extracts interaction.



5. Investigation of transcriptional factors in rat Smad1 intron1 region

5.1 Characterization of rat Smad1 intron1 region

From the analyses of transcriptional activity of rat Smad1 gene, another promoter region was identified in intron1 sequence, as well as several regulatory elements: a positive regulatory element located between I-273 and I-389, and two elements, I-610/490 and I-490/-389, which had different regulatory functions in CFSC-8B cells than in Chang liver cells. Further analysis was performed to characterize these elements. Several putative binding sites for transcriptional factors were predicted after analyzing the proximal intron1 region.

Since transcriptional activity was detected in I-273/+56Smad1p.luc, I-273/+1 sequence was analyzed using the online tool in TRANSFAC database (Fig.36) to search the promoter sequence. However, no classic promoters, such as TATA-box, Inr or CG-rich sequence were found within this region. Putative transcriptional factors for I-273/+1 sequence were listed in Table8.

Three sequences including I-389/-273 (Fig.37), I-490/-389 (Fig.38) and I-610/-490 (Fig.39) were also analyzed and possible transcriptional factors for these regions were listed in Table9. Fewer transcriptional factor binding sites were found in I-389/-273 and I-610/-490 compared with sequence I-490/-389. In sequence I-490/-389, several binding sites for transcriptional factor Oct-1 were detected, as well as one GATA-1 binding motif and one E2F binding motif.

Searching was restricted in vertebrate and the results with a higher score than 85 were displayed.

Database: TRANSFAC MATRIX TABLE, Rel.3.3 06-01-1998
 Query: I-273/+1 (273 bases)
 Taxonomy: Vertebrate
 Threshold: 85.0 point

TFMATRIX entries with High-scoring:

```

1 CTCCAAGCCC AAATACCAGT CGGTGCGAGT GGGTGCTCTG CCTTTAACTG entry      score
51 GTGAACITTA GGAGAACGTA GCTCAGTTCC TGGTGTCACA GGACACAAAT entry      score
      <-----> M00100 CdxA 91.0
      <-----> M00141 Lyf-1 89.6
      <-----> M00100 CdxA 85.9

101 GTTCATGAAC TCAGCATGGG GTTGCTCTCT AATGAAGTCA CCGAATTGGT entry      score
      <-----> M00271 AML-1a 87.4
      <-----> M00041 CRE-BP 87.2
      <-----> M00173 AP-1 86.6
      <-----> M00137 Oct-1 86.5
      <-----> M00041 CRE-BP 86.2
      <-----> M00039 CREB 86.1
      <-----> M00039 CREB 85.6
      <-----> M00113 CREB 85.2

151 CTAAGCCAAG CGTGGACCTC TACAGGGTCA GACTTAACTA GTTAGTACCT entry      score
      <-----> M00235 AhR/Ar 87.6
      <-----> M00240 Nkx-2. 86.0

201 GGCACGCACT AATTAGCTAA TTGTGACGCA GTTAGAGTAA TTTTAAATCA entry      score
      <-----> M00099 S8 95.6
      <-----> M00099 S8 90.4
      <-----> M00133 Tst-1 89.6
      <-----> M00100 CdxA 88.5
      <-----> M00099 S8 88.0
      <-----> M00227 v-Myb 87.7
      <-----> M00235 AhR/Ar 87.6
      <-----> M00173 AP-1 85.6

251 GCTTCTTGTC TTTTCTTTG CAG entry      score
      <-----> M00148 SRY 90.0
      <-----> M00133 Tst-1 89.6
      <-----> M00079 Evi-1 88.6

```

Total 22 high-scoring sites found.
 Max score: 95.6 point, Min score: 85.2 point

Fig.36. Sequence analysis of I-273/+1 of rat Smad1 intron1 region.

Table8. Potential transcriptional factors for I-273/+1 sequence

sequence	Potential transcription factors	
	<i>Factors</i>	<i>Type</i>
I-273/+1	CdxA	Caudal-type homeodomain protein
	LyF-1	For lymphocyte-specific genes
	ALM-1a	Runt-related transcription factor
	CRE-BP	cAMP response element binding protein
	AP-1	activated protein-1
	Oct-1	Octamer-binding factor1
	CREB	CRE-binding protein
	AhR/Ar	Aryl hydrocarbon receptor nuclear translocator
	Nkx-2	Cardiac-specific homeodomain protein
	S8	homeodomain protein
	Tst-1	POU-factor Tst-1/Oct-6
	V-Myb	Avian myoblastosis virus oncogene
	SRY	sex determining region Y gene product
	Evi	ectopic viral integration site1 encoded factor

Database: TRANSFAC MATRIX TABLE, Rel.3.3 06-01-1998
 Query: I-389/-273 (116 bases)
 Taxonomy: Vertebrate
 Threshold: 85.0 point

TFMATRIX entries with High-scoring:

```

1 GCITGGGTGC TGITGGGAGC AATCTATCC CTCCTCTTG TAAAGCTGCC entry      score
                                <-----
                                <-- MO0075 GATA-1 87.3
                                MO0106 CDP CR 86.3
                                <-- MO0076 GATA-2 85.4

51 AATCCTGTTG CCTCAAGCCC CACCCACCCT TATAATCTAA CTTAACCTTA entry      score
    -----
                                <-----
                                MO0075 GATA-1 87.3
                                MO0101 CdxA  86.4
                                MO0148 SRY   85.5
                                MO0076 GATA-2 85.4
                                ----->

101 GTGAGTTACT TAGAGG entry      score

```

Total 5 high-scoring sites found.
 Max score: 87.3 point, Min score: 85.4 point

Fig.37. Sequence analysis of I-389/-273 of rat Smad1 intron1 region.

Database: TRANSFAC MATRIX TABLE, Rel.3.3 06-01-1998
 Query: I-610/-490 (121 bases)
 Taxonomy: Vertebrate
 Threshold: 85.0 point

TFMATRIX entries with High-scoring:

```

1 TTTGGTCGCA TTTTACGCT GTAGTCTTCT GTACCTGTCG GCTCCTGTGT entry      score
                                <-----
                                MO0101 CdxA  85.0

51 ATCAGTCCAT CTCCTCCAG GAGTGTGGT TGCTGGAGTG GTTAGCTTC entry      score
    <-----
                                ----->
                                MO0075 GATA-1 88.6
                                MO0033 p300  86.1
                                MO0271 AML-1a 85.4
                                ----->

101 TTTGCAGTTT TGTGCTTGGT A entry      score

```

Total 4 high-scoring sites found.
 Max score: 88.6 point, Min score: 85.0 point

Fig.38. Sequence analysis of I-610/-490 of rat Smad1 intron1 region.

Table9. Potential transcriptional factors for sequences of rat Smad1 intron1 region

Sequence	Potential transcription factor	
	<i>Factors</i>	<i>Type</i>
I-389/-273	GATAs	globin transcription factors
	CDP CR	cut-like homeodomain protein
	CdxA	Caudal-type homeodomain protein
	SRY	Sex determining region Y gene product
I-490/-389	Oct-1	Octamer-binding factor1
	AML-1a	Runt-related transcription factor
	SRY	sex determining factor
	GATA-1	globin transcription factor-1
	CdxA	Caudal-type homeodomain protein
	COUP-T	Chicken Ovalbumin Upstream Promoter-Transcription factor
	Nkx-2.	Cardiac-specific homeodomain protein
	E2F	control of cell cycle progression and the action of tumor suppressor protein
	ROR α p	RAR-related orphan receptor alpha (steroid hormone nuclear receptor family related to retinoic acid receptors)
I-610/-490	CdxA	Caudal-type homeodomain protein
	GATA-1	globin transcription factor-1
	p300	coactivator
	AML-1a	Runt-related transcription factor

5.2 Interactions between trans-factors and cis-elements in intron1 region

The presence of transcriptional factors interacting with the intron1 region was first addressed by EMSA using nuclear extracts from Chang liver cells and HSC-T6 cells (hepatic stellate cells-T6) with PCR generated probes (Fig.40), Probe3 (I-389/-273), Probe4 (I-490/-389) and Probe5 (I-610/-490). Probe information was listed in Table7 and Table8. For each EMSA reaction, 1ng of radiolabeled probes were incubated with 5µg of nuclear proteins. Results were shown in Fig.41.

No obvious retarded band could be seen in the interaction of Probe5 and nuclear proteins. Using probe3, one major band was found with nuclear extracts from both types of cells, and three other bands were seen with Chang liver nuclear extracts. Different binding patterns also displayed in Probe4/nuclear extracts interaction. Two shift bands were found with Chang liver nuclear extracts and one of them also showed with nuclear extracts from HSC-T6 cells. The difference in shift pattern indicated their different profiles of transcriptional factors between two types of cells.

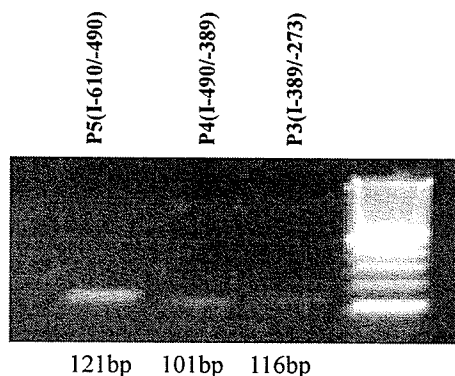


Fig.40. PCR cloning of EMSA probe3, Probe4 and Probe5. Probe3 (I-389/-273), Probe4 (I-490/-389) and Probe5 (I-610/-490) were PCR cloned using the primers listing in Table7. Products were subsequently gel-purified.

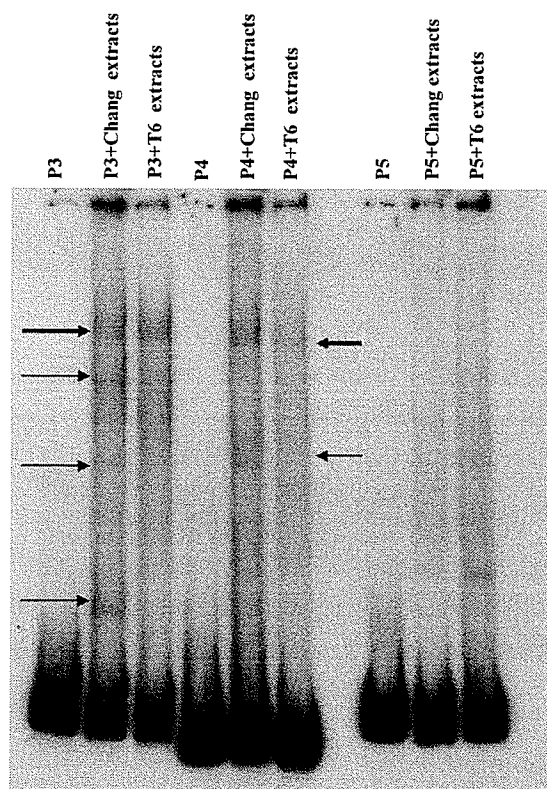


Fig.41. The interaction between nuclear proteins and Probes of rat Smad1 intron1 region. EMSA was carried out with three probes and nuclear proteins isolated from Chang liver cells and HSC-T6 cells. Different binding patterns were seen with nuclear extracts from different cells for both P3 (I-389/-273) and P4 (I-490/389). No clear retarded band was detected with P5 (I-610/-490).

5.3 Determination of trans-factors in intron1 region

Several transcriptional factor binding sites for sequence of I-490/-389 and I-389/-273 were predicted. Competitive binding assay was performed in order to identify the trans-factors. For DNA/protein binding reaction, Probe3 (I-389/-273) and Probe4 (I-490/-389) were incubated with nuclear extracts from Chang liver cells and HSC-T6 cells. Consensus sequences of putative transcriptional factors GATA, E2F and Oct-1 were purchased from Santa Cruz Biotechnology Inc (Table10). In each competitive assay, 100- or 250-fold molar excess amount of unlabeled consensus sequences was added to the binding reaction.

For P3, 100- or 250- molar excess of cold GATA consensus sequence was used in the binding reaction (Fig.42-A). However, no competition was observed with either Chang liver extracts or HSC-T6 extracts. The result suggested that the transcriptional factor binding to I-389/-273 sequence was not GATA.

As shown in Figure 42-B, 100- or 250-fold molar excess of cold GATA, E2F and Oct-1 consensus sequences were used respectively in order to determine the trans-factors for P4. Competitors of GATA and E2F could not abolish the resulting DNA/protein complexes. However, using labeled P4, a gradually decreasing intensity of the shift band was observed when the reaction contained 100-fold to 250-fold molar excess of cold Oct-1 consensus sequence. A slight higher excess of cold sequence was required to thoroughly compete the shift band. Same phenomenon was observed with nuclear extracts from both types of the cells. These results might indicate that sequence I-490/-389 involved the binding site for transcriptional factor Oct-1 and Oct-1 bound to this sequence in both Chang liver cells and HSC-T6 cells.

Table10. Consensus sequences of putative transcription factors in Intron1 region

Name	Consensus Sequences	Length
GATA	5'-CAC TTG ATA ACA GAA AGT GAT AAC TCT-3' 3'-GTG AAC TAT TGT CTT TCA CTA TTG AGA-5'	27bp
E2F	5'-ATT TAA GTT TCG CGC CCT TTC TCA A-3' 3'-TAA ATT CAA AGC GCG GGA AAG AGT T-5'	25bp
Oct-1	5'-TGT CGA ATG CAA ATC ACT AGA A-3' 3'-ACA GCT TAC GTT TAG TGA TCT T-5'	22bp

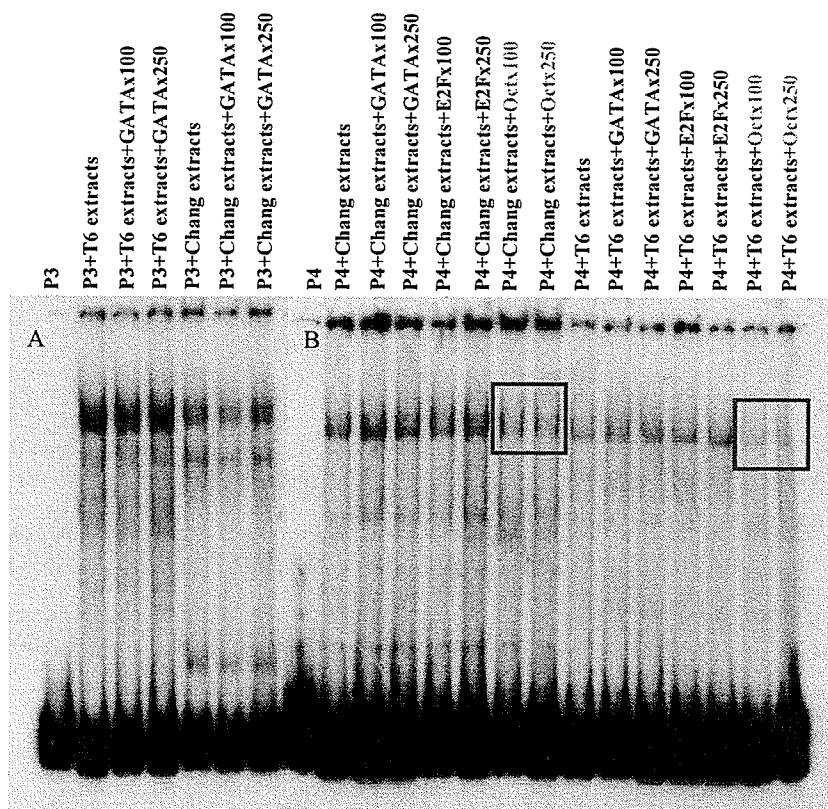


Fig.42. Competition assay using GATA, E2F and Oct-1 consensus sequences. In each reaction, 100- or 250-fold molar excess of different cold consensus sequences were used to compete with labeled Probe3 (1-389/-273) or Probe4 (-490/-389). Nuclear extracts were isolated from Chang liver cells and HSC-T6 cells. A. Competition assay for Probe3. Cold GATA sequence could not abolish the retarded band of P3/nuclear extract complex. B. Competition assay for Probe4. No competition showed using cold GATA and E2F consensus sequences, but cold Oct-1 sequence was found to be able to abolish the shift band. The fading bands were highlighted by green squares.

IV. DISCUSSIONS

Liver fibrosis is a wound healing process of the liver. Due to various insults, hepatocyte damage induces the release of cytokines from Kupffer cells. These cytokines and insults lead to activation of hepatic stellate cells, which synthesize large amounts of extracellular matrix components. In chronic liver injury, accumulation of ECM and damage of hepatocytes cause the disruption of hepatic architecture and metabolism and eventually lead to cirrhosis. It is now believed that the central event of hepatic fibrogenesis is HSC activation (Prosser, Yen et al. 2006).

TGF- β represents a key mediator of fibrogenic process in the liver. TGF- β can not only trigger quiescent HSCs transdifferentiate to myofibroblast-like cells, but also induce the expression of many ECM proteins. In addition, recent research demonstrates that TGF- β causes hepatocytes to undergo apoptosis, therefore facilitating HSC proliferation and ECM accumulation (Breitkopf, Godoy et al. 2006).

Smad proteins are discovered as intracellular mediators transmitting signals of members of TGF- β superfamily. BMP intracellular signaling is usually mediated by Smad1, Smad5 and Smad8, while TGF- β signaling is mediated by Smad2 and Smad3. However it is recognized that type I TGF- β receptor contains certain kinase that could also activate Smad1 and Smad5. As intracellular signaling molecules of TGF- β , the expression and function of Smad2, Smad3, Smad4 and Smad7 have been extensively investigated in animal model of liver fibrosis. On the other hand, the role Smad1 received little attention in study of hepatic fibrosis until recently. Our previous studies demonstrated that the expression of Smad1 was increased during HSC activation and the enhanced expression of Smad1 could be inhibited by TGF- β 1 (Shen, Huang et al. 2003).

These findings reveal a potential role of Smad1 in liver fibrogenesis and emphasize the need to examine mechanism of Smad1 gene regulation in hepatic stellate cells. In addition, there is very little information regarding Smad1 promoter in general. In rat, Smad1 cDNA was cloned and the exon/intron boundary was determined in 2002 (Strausberg, Feingold et al. 2002). However, 5'-flanking region of Smad1 gene has yet been identified.

To understand regulatory mechanism of Smad1 gene expression, we cloned rat Smad1 5'-flanking region and intron1 region, and the transcriptional activities of these regions were examined.

1. 5'-Flanking Promoter Region

1.1 CG-rich promoter

5'-flanking region of rat Smad1 gene (up to 1005bp upstream of exon1) was cloned and transcriptional activity of this sequence was examined in the present study. After examining the sequence by promoter finding software on the internet (<http://molsun1.cbrc.aist.go.jp/research/db/TFSEARCH.html>) no typical TATA-box or Inr was found in this sequence. Moreover, with a DNA probe located at F-163/-56 before transcriptional start site, no DNA/protein interaction was detected, which is consistent with the reports that documented difficulty in detecting nuclear factor binding within that region at TATA-less promoters (Pugh and Tjian 1991; Tyree, George et al. 1993). However, sequence of F-163/+1 contains a CG-rich region including two GC-boxes located at F-96bp and F-52bp respectively. The GC-box is DNA sequence recognized by Sp1 nuclear factor. This information indicates that rat Smad1 has a CG-rich promoter at the 5'-flanking region. CG-rich promoter is primarily found in housekeeping genes (Tamura and Mikoshiba 1991).

Sp1 was the first cloned mammalian transcription factor (Kadonaga, Carner et al. 1987). It binds to CG-rich sequence such as GC-boxes (GGGCGG) and GT-boxes (CACCC). Sp1 was identified in the studies of the transcription of SV40 gene. It was found that Sp1 could stimulate the expression of SV40 large T antigen and the stimulation was dependent on presence of GC-box in the promoter region. Due to its specific binding to GC-box, Sp1 is named as Specificity Protein1.

Early studies indicated that Sp1 was responsible for initiation of transcription at TATA-less promoters (Lu, Lee et al. 1994). This finding leads to the notion that Sp1 acts as a basal transcriptional factor and Sp1 binding sites represent constitutive elements that support basal transcription at CG-rich promoters (Black, Black et al. 2001). Sp1 is subject to extensive post-translational modification, which indicates that activity of Sp1 is likely to be regulated at post-translational level (Jackson, MacDonald et al. 1990).

At present, it is known that Sp1 belongs to Sp/KLF (Sp1/Kruppel-Like Factor) family, which contains 20 identified members. Sp/KLF members are involved in many growth-related signaling pathways and they have great effects on cell proliferation. Other than growth control, many factors in this family involve in some aspects of tumorigenesis (Turner and Crossley 1999). In activated HSCs, at least three members have been identified including Sp1, BTEB1 and KLF6 (formerly Zf9). These factors have the ability to regulate transcription of collagen I (Mann and Smart 2002). The mutagenesis of the Sp1 binding site on the promoter results in greatly reduced transcriptional activity of collagen I. Moreover, the interaction between Sp1 and Smad3 is essential for basal and TGF- β stimulated expression of collagen 1 in HSCs. Furthermore, Sp1 expression and DNA binding activity are enhanced during HSC activation by a post-transcriptional mechanism (Rippe, Almounajed et al. 1995).

1.2 C/EBP element

Luciferase assay also identified a positive regulatory element between F-163bp and F-56bp and transcriptional consensus binding site search predicted several potential nuclear transcriptional factor binding sites in this region. One of the transcriptional binding sites is the sequence for nuclear transcriptional factor C/EBP, which located at F-162 of the rat Smad1 promoter.

C/EBP belongs to a family of transcriptional factors that bind to CCAAT-box (Lekstrom-Himes and Xanthopoulos 1998). C/EBP consists of an activation domain, DNA binding basic region and a leucine-rich dimerization domain. This family includes six members and the expression of these factors is quite unique. The expression of C/EBP γ and C/EBP ζ is ubiquitous in the tissue while other C/EBPs are expressed in a tissue-specific manner. Three isoforms of C/EBPs were found in the liver, which are C/EBP α , C/EBP β and C/EBP δ . C/EBP members play pivot role in regulation of cell differentiation, inflammatory response, liver regeneration and numerous other cellular responses (Buck and Chojkier 2003).

C/EBP binding site was also found in collagen I promoter. It is documented that TGF- β stimulation of collagen I expression was in part due to the increased expression of C/EBP β . It was demonstrated that TGF- β stimulated the formation of H₂O₂ in HSCs and accumulation of intracellular H₂O₂ stimulates the expression of C/EBP β . Therefore, increased level of C/EBP β was able to bind to the CCAAT-box on collagen I promoter and increased the transcription rate of collagen I (Greenwel, Dominguez-Rosales et al. 2000).

2. Intron1 Region

2.1 Two transcripts of rat Smad1

There are two transcripts of rat Smad1 in GenBank with the length of 1425bp (Genbank No. AF067727) and 3010bp (Genbank No. NM_013130) respectively. The differences between these two transcripts are at the both ends. There is a lack of 320bp non-coding sequence at the 5'-end and 1227bp non-coding sequence at the 3'-end. The genomic sequence of rat Smad1 published in Ensembl website in 2004 was organized according to the shorter transcript of 1425bp, but in 2005, the genomic sequence was updated based on the transcript of 3010bp. Compared with the transcript of 3010bp, the transcript of 1425bp misses a 157bp of exon1 and 163bp of exon2 at the 5'-end, while at the 3'-end, there is a big truncate of exon7 about 1227bp (Fig.8 and 10). There are several possibilities for this phenomenon. One of the possibilities is that there is a promoter in the intron1 region and Smad1 transcription may initiate at different position according to different tissues or cell types. Therefore, in order to understanding whether there is a promoter in the intron1 region, we cloned the intron1 region and examined the transcriptional activity in this region.

2.2 Transcriptional activity of the intron1 region

A 2015bp of rat Smad1 intron1 region was cloned and analyzed for the transcriptional activity. It was found that there is a promote-like activity in the intron1 region. Compared to the 5'-flanking region, the promoter activity of the intron1 region was much lower and there is no TATA-box, Inr or CG-rich sequence in the intron1 region.

Although the promoter activity is low for the intron1 region, the activity may be greater in other cell types than hepatocyte. It is known that several genes contain multiple promoters in their gene structure, such as A1 adenosine receptor, acetyl-CoA synthetase, acetylcholinesterase gene and growth hormone-releasing hormone (Lee,

Minowa et al. 1996). Some of the promoters were utilized in a tissue specific manner to express different products in different tissues and some of them activated differentially during ontogeny. In current study, both the 5'-flanking region and the intron1 region contain promoter activity with different strength. Since both transcripts will result a same coding protein, the two different transcription start sites may indicate a mechanism that makes sure to express this important gene effectively in different tissues or cell types.

2.3 *Oct-1 binding site*

Some regulatory elements in the intron1 region were found by deletional analyses and the putative binding factors were predicted using consensus site finding tool on TRBSFAC website. To confirm the putative binding factors, nuclear protein competition assay was carried out. A putative nuclear factor binding site for Oct-1 was found between I-490bp and I-389bp region.

Octamer-binding factor-1 (Oct-1 or OTF-1) is a transcriptional factor of POU family (Latchman 1999). Members of the POU (Pit-Oct-Unc) family were originally identified on the basis of a common DNA binding domain in the mammalian factors Pit-1, Oct-1 as well as the nematode protein Unc-86. Octamer-binding factors, which recognize the sequence ATGCAAAT, are involved in a broad range of biological processes. Oct-1 is expressed in all cell types and plays an important role in regulation of housekeeping gene expression and development of immune response (Remenyi, Tomilin et al. 2002). However, Oct-1 regulation of Smad1 gene expression has not been defined.

3. Transcriptional activity of Smad1 gene in HSCs

Hepatic stellate cells have two distinct states (Saile, Knittel et al. 1997). In the normal liver, they are quiescent and replete with retinoid droplets. With liver injury, HSCs activate and transform into myofibroblast-like cells. In activated state, HSCs loss of the retinol stores and become to enlarged cells expressing α -smooth muscle actin (α -SMA). The activation of HSCs is considered to be the central event of hepatic fibrogenesis.

Our previous finding demonstrated that Smad1 expression increased during HSC activation, which leads to the hypothesis that transcriptional activity of rat Smad1 increases during HSC activation. To compare the transcriptional activity of rat Smad1 in different states of HSCs, the 5'-flanking region and intron1 region were transfected into quiescent or activated CFSC (cirrhotic fat-storing cell)-8B cells. The transcriptional activities of these two regions were very strong in activated CFSC-8B cells. However, in quiescent cells, the activities were undetectable for both regions. It seemed that the function of rat Smad1 promoters was 'silenced' in quiescent HSCs. This finding confirms our hypothesis that transcriptional activity of rat Smad1 increases during HSC activation. The change of transcriptional activity of rat Smad1 in activated HSCs could be resulted from enhanced binding of transcriptional factors, which may play an important role in the regulation of Smad1 gene expression during the activation of HSCs. However, the exact mechanism still needs to be further investigated.

4. Smad transcriptional regulation

There were very few reports regarding the transcriptional regulation of Smad proteins although the promoters of Smad4, Smad6 and Smad7 have been

characterized. Human Smad4 (Minami, Kitazawa et al. 1998) and rat Smad7 (Stopa, Benes et al. 2000) contained two different types of TATA-like sequences ('TAAAAT' and 'TTTAAA') to initiate transcription respectively. However, a CG-rich promoter with multiple Sp1 binding sites was identified in mouse Smad6 gene (Watanabe, Masuyama et al. 2000). At present, there is no report regarding the promoter or 5'-flanking region of Smad1, therefore, our finding indicates the first report about Smad1 promoter and the 5'-flanking region.

5. Future direction

In the present study, we found that rat Smad1 gene had two promoter regions in its 5'-flanking region and intron1 region respectively. Many literatures have revealed that some of the gene can have more than one promoter and these promoters can be used at different time or in different situations. Therefore, it will be a very interesting question to be answered that how rat Smad1 uses the promoters for its gene expression.

We've also found that promoter activity of rat Smad1 gene was 'silenced' in quiescent HSCs and the expression of rat Smad1 gene was increased in activated HSCs. Therefore, our future study can focus on what 'opens' rat Smad1 promoter activity during HSC activation. HSC activation requires global reprogramming of gene expression. At transcriptional level, there should be some key transcription factors to control this program. To discover which factor 'open' rat Smad1 gene expression can help us to find the 'controller' of this cellular event on HSC activation.

Conclusions:

- We have cloned parts of the 5'-flanking region and the intron1 region of rat Smad1 gene. The 5'-flanking promoter is a CG-rich promoter with two Sp1 binding sites. The intron1 promoter contains an Oct-1 binding site between I-490bp and I-389bp.
- The transcriptional activity of the 5'-flanking promoter was higher than that of the intron1 promoter indicating that Smad1 gene expression usually initiates from the 5'-flanking promoter.
- Transcriptional activities of both the 5'-flanking promoter and the intron1 promoter were increased in the activated HSCs compared to that in the quiescent HSCs indicating an increase in Smad1 gene expression during the activation of HSCs.

V. REFERENCES

- Afdhal, N. H. and D. Nunes (2004). "Evaluation of liver fibrosis: a concise review." Am J Gastroenterol **99**(6): 1160-74.
- Albanis, E. and S. L. Friedman (2001). "Hepatic fibrosis. Pathogenesis and principles of therapy." Clin Liver Dis **5**(2): 315-34, v-vi.
- Alcolado, R., M. J. Arthur, et al. (1997). "Pathogenesis of liver fibrosis." Clin Sci (Lond) **92**(2): 103-12.
- Arii, S. and M. Imamura (2000). "Physiological role of sinusoidal endothelial cells and Kupffer cells and their implication in the pathogenesis of liver injury." J Hepatobiliary Pancreat Surg **7**(1): 40-8.
- Attisano, L. and S. Tuen Lee-Hoeflich (2001). "The Smads." Genome Biol **2**(8): REVIEWS3010.
- Attisano, L. and J. L. Wrana (1998). "Mads and Smads in TGF beta signalling." Curr Opin Cell Biol **10**(2): 188-94.
- Bachem, M. G., K. M. Sell, et al. (1993). "Tumor necrosis factor alpha (TNF alpha) and transforming growth factor beta 1 (TGF beta 1) stimulate fibronectin synthesis and the transdifferentiation of fat-storing cells in the rat liver into myofibroblasts." Virchows Arch B Cell Pathol Incl Mol Pathol **63**(2): 123-30.
- Barcellos-Hoff, M. H. (1996). "Latency and activation in the control of TGF-beta." J Mammary Gland Biol Neoplasia **1**(4): 353-63.
- Bataller, R. and D. A. Brenner (2001). "Hepatic stellate cells as a target for the treatment of liver fibrosis." Semin Liver Dis **21**(3): 437-51.
- Bataller, R. and D. A. Brenner (2005). "Liver fibrosis." J Clin Invest **115**(2): 209-18.
- Benoist, C. and P. Chambon (1981). "In vivo sequence requirements of the SV40 early promotor region." Nature **290**(5804): 304-10.

- Bissell, D. M., S. S. Wang, et al. (1995). "Cell-specific expression of transforming growth factor-beta in rat liver. Evidence for autocrine regulation of hepatocyte proliferation." J Clin Invest **96**(1): 447-55.
- Black, A. R., J. D. Black, et al. (2001). "Sp1 and kruppel-like factor family of transcription factors in cell growth regulation and cancer." J Cell Physiol **188**(2): 143-60.
- Bouwens, L., P. De Bleser, et al. (1992). "Liver cell heterogeneity: functions of non-parenchymal cells." Enzyme **46**(1-3): 155-68.
- Breitkopf, K., S. Haas, et al. (2005). "Anti-TGF-beta strategies for the treatment of chronic liver disease." Alcohol Clin Exp Res **29**(11 Suppl): 121S-131S.
- Breitkopf, K., P. Godoy, et al. (2006). "TGF-beta/Smad signaling in the injured liver." Z Gastroenterol **44**(1): 57-66.
- Brown, D. D. (1982). "Gene expression in eukaryotes." Oncodev Biol Med **4**(1-2): 9-29.
- Brunt, E. M. (2004). "Nonalcoholic steatohepatitis." Semin Liver Dis **24**(1): 3-20.
- Buck, M. and M. Chojkier (2003). "Signal transduction in the liver: C/EBPbeta modulates cell proliferation and survival." Hepatology **37**(4): 731-8.
- Burt, A. D., M. R. Griffiths, et al. (1990). "Ultrastructural localization of extracellular matrix proteins in liver biopsies using ultracryomicrotomy and immuno-gold labelling." Histopathology **16**(1): 53-8.
- Casini, A., M. Pinzani, et al. (1993). "Regulation of extracellular matrix synthesis by transforming growth factor beta 1 in human fat-storing cells." Gastroenterology **105**(1): 245-53.
- Castilla, A., J. Prieto, et al. (1991). "Transforming growth factors beta 1 and alpha in chronic liver disease. Effects of interferon alfa therapy." N Engl J Med

324(14): 933-40.

Catala, M., A. Anton, et al. (1999). "Characterization of the simultaneous binding of Escherichia coli endotoxin to Kupffer and endothelial liver cells by flow cytometry." Cytometry **36**(2): 123-30.

Cheifetz, S., J. L. Andres, et al. (1988). "The transforming growth factor-beta receptor type III is a membrane proteoglycan. Domain structure of the receptor." J Biol Chem **263**(32): 16984-91.

Chen, Y., J. J. Lebrun, et al. (1996). "Regulation of transforming growth factor beta- and activin-induced transcription by mammalian Mad proteins." Proc Natl Acad Sci U S A **93**(23): 12992-7.

Cooper, S. (1981). "The central dogma of cell biology." Cell Biol Int Rep **5**(6): 539-49.

Davis, G. L., J. E. Albright, et al. (2003). "Projecting future complications of chronic hepatitis C in the United States." Liver Transpl **9**(4): 331-8.

de Bleser, P. J., P. Jannes, et al. (1995). "Insulinlike growth factor-II/mannose 6-phosphate receptor is expressed on CCl4-exposed rat fat-storing cells and facilitates activation of latent transforming growth factor-beta in cocultures with sinusoidal endothelial cells." Hepatology **21**(5): 1429-37.

Derynck, R. and Y. E. Zhang (2003). "Smad-dependent and Smad-independent pathways in TGF-beta family signalling." Nature **425**(6958): 577-84.

Docherty, A. J., A. Lyons, et al. (1985). "Sequence of human tissue inhibitor of metalloproteinases and its identity to erythroid-potentiating activity." Nature **318**(6041): 66-9.

Dooley, S., B. Delvoux, et al. (2000). "Modulation of transforming growth factor beta response and signaling during transdifferentiation of rat hepatic stellate cells to

- myofibroblasts." Hepatology **31**(5): 1094-106.
- Dooley, S., J. Hamzavi, et al. (2003). "Smad7 prevents activation of hepatic stellate cells and liver fibrosis in rats." Gastroenterology **125**(1): 178-91.
- Edwards, D. R., G. Murphy, et al. (1987). "Transforming growth factor beta modulates the expression of collagenase and metalloproteinase inhibitor." Embo J **6**(7): 1899-904.
- Etienne, J., D. Brault, et al. (1990). "["Cis" and "trans" regulator elements of transcription]." Ann Biol Clin (Paris) **48**(10): 681-94.
- Fallowfield, J. A. and J. P. Iredale (2004). "Targeted treatments for cirrhosis." Expert Opin Ther Targets **8**(5): 423-35.
- Feng, X. H., Y. Zhang, et al. (1998). "The tumor suppressor Smad4/DPC4 and transcriptional adaptor CBP/p300 are coactivators for smad3 in TGF-beta-induced transcriptional activation." Genes Dev **12**(14): 2153-63.
- Friedman, S. L. (1990). "Cellular sources of collagen and regulation of collagen production in liver." Semin Liver Dis **10**(1): 20-9.
- Friedman, S. L. (2000). "Molecular regulation of hepatic fibrosis, an integrated cellular response to tissue injury." J Biol Chem **275**(4): 2247-50.
- Friedman, S. L. (2003). "Liver fibrosis -- from bench to bedside." J Hepatol **38 Suppl 1**: S38-53.
- Friedman, S. L., D. C. Rockey, et al. (1992). "Isolated hepatic lipocytes and Kupffer cells from normal human liver: morphological and functional characteristics in primary culture." Hepatology **15**(2): 234-43.
- Friedman, S. L., F. J. Roll, et al. (1989). "Maintenance of differentiated phenotype of cultured rat hepatic lipocytes by basement membrane matrix." J Biol Chem **264**(18): 10756-62.

- Friedman, S. L., F. J. Roll, et al. (1985). "Hepatic lipocytes: the principal collagen-producing cells of normal rat liver." Proc Natl Acad Sci U S A **82**(24): 8681-5.
- Friedman, S. L., G. Yamasaki, et al. (1994). "Modulation of transforming growth factor beta receptors of rat lipocytes during the hepatic wound healing response. Enhanced binding and reduced gene expression accompany cellular activation in culture and in vivo." J Biol Chem **269**(14): 10551-8.
- Geerts, A. (2001). "History, heterogeneity, developmental biology, and functions of quiescent hepatic stellate cells." Semin Liver Dis **21**(3): 311-35.
- Geerts, A., R. Vrijssen, et al. (1989). "In vitro differentiation of fat-storing cells parallels marked increase of collagen synthesis and secretion." J Hepatol **9**(1): 59-68.
- Gines, P., A. Cardenas, et al. (2004). "Management of cirrhosis and ascites." N Engl J Med **350**(16): 1646-54.
- Greenwel, P., J. A. Dominguez-Rosales, et al. (2000). "Hydrogen peroxide: a link between acetaldehyde-elicited alpha1(I) collagen gene up-regulation and oxidative stress in mouse hepatic stellate cells." Hepatology **31**(1): 109-16.
- Gressner, A. M. and M. G. Bachem (1990). "Cellular sources of noncollagenous matrix proteins: role of fat-storing cells in fibrogenesis." Semin Liver Dis **10**(1): 30-46.
- Gressner, A. M., R. Weiskirchen, et al. (2002). "Roles of TGF-beta in hepatic fibrosis." Front Biosci **7**: d793-807.
- Hammel, P., A. Couvelard, et al. (2001). "Regression of liver fibrosis after biliary drainage in patients with chronic pancreatitis and stenosis of the common bile duct." N Engl J Med **344**(6): 418-23.

- Hata, A., J. Seoane, et al. (2000). "OAZ uses distinct DNA- and protein-binding zinc fingers in separate BMP-Smad and Olf signaling pathways." Cell **100**(2): 229-40.
- Hayashi, H., S. Abdollah, et al. (1997). "The MAD-related protein Smad7 associates with the TGFbeta receptor and functions as an antagonist of TGFbeta signaling." Cell **89**(7): 1165-73.
- Hendriks, H. F., W. S. Blaner, et al. (1988). "Distributions of retinoids, retinoid-binding proteins and related parameters in different types of liver cells isolated from young and old rats." Eur J Biochem **171**(1-2): 237-44.
- Hsuan, J. J. (1989). "Transforming growth factors beta." Br Med Bull **45**(2): 425-37.
- Imai, S., M. Okuno, et al. (1997). "9,13-di-cis-Retinoic acid induces the production of tPA and activation of latent TGF-beta via RAR alpha in a human liver stellate cell line, LI90." FEBS Lett **411**(1): 102-6.
- Ishisaki, A., K. Yamato, et al. (1999). "Differential inhibition of Smad6 and Smad7 on bone morphogenetic protein- and activin-mediated growth arrest and apoptosis in B cells." J Biol Chem **274**(19): 13637-42.
- Itoh, F., H. Asao, et al. (2001). "Promoting bone morphogenetic protein signaling through negative regulation of inhibitory Smads." Embo J **20**(15): 4132-42.
- Itoh, S., F. Itoh, et al. (2000). "Signaling of transforming growth factor-beta family members through Smad proteins." Eur J Biochem **267**(24): 6954-67.
- Jakowlew, S. B., J. E. Mead, et al. (1991). "Transforming growth factor-beta (TGF-beta) isoforms in rat liver regeneration: messenger RNA expression and activation of latent TGF-beta." Cell Regul **2**(7): 535-48.
- Jackson, S. P., J. J. MacDonald, et al. (1990). "GC box binding induces phosphorylation of Sp1 by a DNA-dependent protein kinase." Cell **63**(1):

155-65.

- Jonk, L. J., S. Itoh, et al. (1998). "Identification and functional characterization of a Smad binding element (SBE) in the JunB promoter that acts as a transforming growth factor-beta, activin, and bone morphogenetic protein-inducible enhancer." J Biol Chem **273**(33): 21145-52.
- Kadonaga, J. T., K. R. Carner, et al. (1987). "Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain." Cell **51**(6): 1079-90.
- Kinnman, N. and C. Housset (2002). "Peribiliary myofibroblasts in biliary type liver fibrosis." Front Biosci **7**: d496-503.
- Kmiec, Z. (2001). "Cooperation of liver cells in health and disease." Adv Anat Embryol Cell Biol **161**: III-XIII, 1-151.
- Kojima, S., S. Hayashi, et al. (2000). "Transcriptional activation of urokinase by the Kruppel-like factor Zf9/COPEB activates latent TGF-beta1 in vascular endothelial cells." Blood **95**(4): 1309-16.
- Kovalszky, I., P. Nagy, et al. (1998). "Experimental and human liver fibrogenesis." Scand J Gastroenterol Suppl **228**: 51-5.
- Latchman, D. S. (1999). "POU family transcription factors in the nervous system." J Cell Physiol **179**(2): 126-33.
- Lee, S. H., M. T. Minowa, et al. (1996). "Two distinct promoters drive transcription of the human D1A dopamine receptor gene." J Biol Chem **271**(41): 25292-9.
- Lekstrom-Himes, J. and K. G. Xanthopoulos (1998). "Biological role of the CCAAT/enhancer-binding protein family of transcription factors." J Biol Chem **273**(44): 28545-8.
- Li, Y. L., M. Sato, et al. (1999). "Regulatory role of extracellular matrix components

in expression of matrix metalloproteinases in cultured hepatic stellate cells." Cell Struct Funct **24**(5): 255-61.

Liu, C., M. D. Gaca, et al. (2003). "Smads 2 and 3 are differentially activated by transforming growth factor-beta (TGF-beta) in quiescent and activated hepatic stellate cells. Constitutive nuclear localization of Smads in activated cells is TGF-beta-independent." J Biol Chem **278**(13): 11721-8.

Liu, F., A. Hata, et al. (1996). "A human Mad protein acting as a BMP-regulated transcriptional activator." Nature **381**(6583): 620-3.

Lo, R. S., Y. G. Chen, et al. (1998). "The L3 loop: a structural motif determining specific interactions between SMAD proteins and TGF-beta receptors." Embo J **17**(4): 996-1005.

Lu, J., W. Lee, et al. (1994). "Start site selection by Sp1 in the TATA-less human Ha-ras promoter." J Biol Chem **269**(7): 5391-402.

Maher, J. J. (2001). "Interactions between hepatic stellate cells and the immune system." Semin Liver Dis **21**(3): 417-26.

Mann, D. A. and D. E. Smart (2002). "Transcriptional regulation of hepatic stellate cell activation." Gut **50**(6): 891-6.

Marek, A., J. Brodzicki, et al. (2002). "TGF-beta (transforming growth factor-beta) in chronic inflammatory conditions - a new diagnostic and prognostic marker?" Med Sci Monit **8**(7): RA145-51.

Massague, J. (1998). "TGF-beta signal transduction." Annu Rev Biochem **67**: 753-91.

Massague, J. (2000). "How cells read TGF-beta signals." Nat Rev Mol Cell Biol **1**(3): 169-78.

Massague, J. and Y. G. Chen (2000). "Controlling TGF-beta signaling." Genes Dev **14**(6): 627-44.

- Minami, R., R. Kitazawa, et al. (1998). "Analysis of 5'-flanking region of human Smad4 (DPC4) gene." Biochim Biophys Acta **1443**(1-2): 182-5.
- Moustakas, A., S. Soucheinytskyi, et al. (2001). "Smad regulation in TGF-beta signal transduction." J Cell Sci **114**(Pt 24): 4359-69.
- Munger, J. S., J. G. Harpel, et al. (1997). "Latent transforming growth factor-beta: structural features and mechanisms of activation." Kidney Int **51**(5): 1376-82.
- Murakami, G., T. Watabe, et al. (2003). "Cooperative inhibition of bone morphogenetic protein signaling by Smurf1 and inhibitory Smads." Mol Biol Cell **14**(7): 2809-17.
- Murphy, G., A. J. Docherty, et al. (1991). "Metalloproteinases and tissue damage." Br J Rheumatol **30 Suppl 1**: 25-31.
- Neubauer, K., B. Saile, et al. (2001). "Liver fibrosis and altered matrix synthesis." Can J Gastroenterol **15**(3): 187-93.
- Newfeld, S. J., E. H. Chartoff, et al. (1996). "Mothers against dpp encodes a conserved cytoplasmic protein required in DPP/TGF-beta responsive cells." Development **122**(7): 2099-108.
- Okuno, M., S. Kojima, et al. (2002). "Retinoids in liver fibrosis and cancer." Front Biosci **7**: d204-18.
- Okuyama, H., Y. Shimahara, et al. (2002). "The hepatic stellate cell in the post-genomic era." Histol Histopathol **17**(2): 487-95.
- Overall, C. M., J. L. Wrana, et al. (1989). "Independent regulation of collagenase, 72-kDa progelatinase, and metalloendoproteinase inhibitor expression in human fibroblasts by transforming growth factor-beta." J Biol Chem **264**(3): 1860-9.
- Pinzani, M. (2002). "PDGF and signal transduction in hepatic stellate cells." Front

- Biosci **7**: d1720-6.
- Popper, H. and S. Uenfriend (1970). "Hepatic fibrosis. Correlation of biochemical and morphologic investigations." Am J Med **49**: 707-21.
- Poynard, T., P. Bedossa, et al. (1997). "Natural history of liver fibrosis progression in patients with chronic hepatitis C. The OBSVIRC, METAVIR, CLINIVIR, and DOSVIRC groups." Lancet **349**(9055): 825-32.
- Poynard, T., V. Ratziu, et al. (2000). "Natural history of HCV infection." Baillieres Best Pract Res Clin Gastroenterol **14**(2): 211-28.
- Prosser, C. C., R. D. Yen, et al. (2006). "Molecular therapy for hepatic injury and fibrosis: where are we?" World J Gastroenterol **12**(4): 509-15.
- Pugh, B. F. and R. Tjian (1991). "Transcription from a TATA-less promoter requires a multisubunit TFIID complex." Genes Dev **5**(11): 1935-45.
- Rafferty, L. A., V. Twombly, et al. (1995). "Genetic screens to identify elements of the decapentaplegic signaling pathway in Drosophila." Genetics **139**(1): 241-54.
- Reeves, H. L. and S. L. Friedman (2002). "Activation of hepatic stellate cells--a key issue in liver fibrosis." Front Biosci **7**: d808-26.
- Remenyi, A., A. Tomilin, et al. (2002). "Differential activity by DNA-induced quarternary structures of POU transcription factors." Biochem Pharmacol **64**(5-6): 979-84.
- Rifkin, D. B., R. Mazzieri, et al. (1999). "Proteolytic control of growth factor availability." Apmis **107**(1): 80-5.
- Riggins, G. J., S. Thiagalingam, et al. (1996). "Mad-related genes in the human." Nat Genet **13**(3): 347-9.
- Rippe, R. A., G. Almounajed, et al. (1995). "Sp1 binding activity increases in activated Ito cells." Hepatology **22**(1): 241-51.

- Rojkind, M., M. A. Giambrone, et al. (1979). "Collagen types in normal and cirrhotic liver." Gastroenterology **76**(4): 710-9.
- Saile, B., T. Knittel, et al. (1997). "CD95/CD95L-mediated apoptosis of the hepatic stellate cell. A mechanism terminating uncontrolled hepatic stellate cell proliferation during hepatic tissue repair." Am J Pathol **151**(5): 1265-72.
- Sarbah, S. A., T. Gramlich, et al. (2004). "Risk factors for hepatocellular carcinoma in patients with cirrhosis." Dig Dis Sci **49**(5): 850-3.
- Sarem, M., R. Znaidak, et al. (2006). "[Hepatic stellate cells: it's role in normal and pathological conditions]." Gastroenterol Hepatol **29**(2): 93-101.
- Sasse, D., U. M. Spornitz, et al. (1992). "Liver architecture." Enzyme **46**(1-3): 8-32.
- Savage, C., P. Das, et al. (1996). "Caenorhabditis elegans genes sma-2, sma-3, and sma-4 define a conserved family of transforming growth factor beta pathway components." Proc Natl Acad Sci U S A **93**(2): 790-4.
- Schnabl, B., Y. O. Kweon, et al. (2001). "The role of Smad3 in mediating mouse hepatic stellate cell activation." Hepatology **34**(1): 89-100.
- Schuppan, D. (1990). "Structure of the extracellular matrix in normal and fibrotic liver: collagens and glycoproteins." Semin Liver Dis **10**(1): 1-10.
- Shah, V. (2001). "Cellular and molecular basis of portal hypertension." Clin Liver Dis **5**(3): 629-44.
- Shames, B. D., C. H. Selzman, et al. (1998). "Genes don't count." Arch Surg **133**(6): 667-9.
- Shen, H., G. Huang, et al. (2003). "Transforming growth factor-beta1 downregulation of Smad1 gene expression in rat hepatic stellate cells." Am J Physiol Gastrointest Liver Physiol **285**(3): G539-46.
- Shenk, T. (1981). "Transcriptional control regions: nucleotide sequence requirements

- p>for initiation by RNA polymerase II and III."
- Curr Top Microbiol Immunol
- 93**
- : 25-46.
- Shi, Y., Y. F. Wang, et al. (1998). "Crystal structure of a Smad MH1 domain bound to DNA: insights on DNA binding in TGF-beta signaling." Cell **94**(5): 585-94.
- Sohara, N., I. Znoyko, et al. (2002). "Reversal of activation of human myofibroblast-like cells by culture on a basement membrane-like substrate." J Hepatol **37**(2): 214-21.
- Soyer, M. T., R. Ceballos, et al. (1976). "Reversibility of severe hepatic damage caused by jejunoileal bypass after re-establishment of normal intestinal continuity." Surgery **79**(5): 601-4.
- Stopa, M., V. Benes, et al. (2000). "Genomic locus and promoter region of rat Smad7, an important antagonist of TGFbeta signaling." Mamm Genome **11**(2): 169-76.
- Strausberg, R. L., E. A. Feingold, et al. (2002). "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." Proc Natl Acad Sci U S A **99**(26): 16899-903.
- Tamura, T. and K. Mikoshiba (1991). "Role of a GC-rich motif in transcription regulation of the adenovirus type 2 IVa2 promoter which lacks typical TATA-box element." FEBS Lett **282**(1): 87-90.
- Tangkijvanich, P., S. P. Tam, et al. (2001). "Wound-induced migration of rat hepatic stellate cells is modulated by endothelin-1 through rho-kinase-mediated alterations in the acto-myosin cytoskeleton." Hepatology **33**(1): 74-80.
- ten Dijke, P., K. Miyazono, et al. (2000). "Signaling inputs converge on nuclear effectors in TGF-beta signaling." Trends Biochem Sci **25**(2): 64-70.
- Travers, A. (1974). "RNA polymerase--promoter interactions: some general

- principles." Cell **3**(2): 97-104.
- Turner, J. and M. Crossley (1999). "Mammalian Kruppel-like transcription factors: more than just a pretty finger." Trends Biochem Sci **24**(6): 236-40.
- Tyree, C. M., C. P. George, et al. (1993). "Identification of a minimal set of proteins that is sufficient for accurate initiation of transcription by RNA polymerase II." Genes Dev **7**(7A): 1254-65.
- Watanabe, M., N. Masuyama, et al. (2000). "Regulation of intracellular dynamics of Smad4 by its leucine-rich nuclear export signal." EMBO Rep **1**(2): 176-82.
- Weis, L. and D. Reinberg (1992). "Transcription by RNA polymerase II: initiator-directed formation of transcription-competent complexes." Faseb J **6**(14): 3300-9.
- Wells, R. G. (2005). "The role of matrix stiffness in hepatic stellate cell activation and liver fibrosis." J Clin Gastroenterol **39**(4 Suppl 2): S158-61.
- Wheeler, M. D. (2003). "Endotoxin and Kupffer cell activation in alcoholic liver disease." Alcohol Res Health **27**(4): 300-6.
- Wu, J. and M. A. Zern (2000). "Hepatic stellate cells: a target for the treatment of liver fibrosis." J Gastroenterol **35**(9): 665-72.
- Wu, R. Y., Y. Zhang, et al. (1997). "Heteromeric and homomeric interactions correlate with signaling activity and functional cooperativity of Smad3 and Smad4/DPC4." Mol Cell Biol **17**(5): 2521-8.
- Xiao, Z., N. Watson, et al. (2001). "Nucleocytoplasmic shuttling of Smad1 conferred by its nuclear localization and nuclear export signals." J Biol Chem **276**(42): 39404-10.
- Xu, X. B., X. S. Leng, et al. (2004). "Inhibitory effect of retroviral vector containing anti-sense Smad4 gene on Ito cell line, LI90." Chin Med J (Engl) **117**(8):

1170-7.

- Yata, Y., P. Gotwals, et al. (2002). "Dose-dependent inhibition of hepatic fibrosis in mice by a TGF-beta soluble receptor: implications for antifibrotic therapy." Hepatology **35**(5): 1022-30.
- Yee, J. A., L. Yan, et al. (1993). "Plasminogen-dependent activation of latent transforming growth factor beta (TGF beta) by growing cultures of osteoblast-like cells." J Cell Physiol **157**(3): 528-34.
- Yue, J., M. T. Hartsough, et al. (1999). "Cloning and expression of a rat Smad1: regulation by TGFbeta and modulation by the Ras/MEK pathway." J Cell Physiol **178**(3): 387-96.
- Zanninelli, G., A. Vetuschi, et al. (2006). "Smad3 knock-out mice as a useful model to study intestinal fibrogenesis." World J Gastroenterol **12**(8): 1211-8.
- Zhang, W., J. Ou, et al. (2000). "Synergistic cooperation between Sp1 and Smad3/Smad4 mediates transforming growth factor beta1 stimulation of alpha 2(I)-collagen (COL1A2) transcription." J Biol Chem **275**(50): 39237-45.
- Zhao, G. Q. and B. L. Hogan (1997). "Evidence that Mothers-against-dpp-related 1 (Madr1) plays a role in the initiation and maintenance of spermatogenesis in the mouse." Mech Dev **61**(1-2): 63-73.
- Zhu, H., P. Kavsak, et al. (1999). "A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic pattern formation." Nature **400**(6745): 687-93.