Aus der I. Medizinische Klinik und Poliklinik
des Zentrums für Innere Medizin
des Universitätsklinikums Hamburg-Eppendorf
Direktor …Prof. Dr. Ansgar W. Lohse

Identification and Characterization of the Promoter Region of the SLA/LP Gene

Dissertation
zur Erlangung des Grades eines Doktors der Medizin der
Medizinischen Fakultät der Universität Hamburg
vorgelegt von

......Chunxia Wang......
aus ....Shandong, P.R.China..........................
Hamburg .... 2006 ....

Angenommen von der Medizinischen Fakultät
der Universität Hamburg am: 19. 12. 2006
Veröffentlicht mit Genehmigung der Medizinischen
Fakultät der Universität Hamburg
Prüfungsausschuss, der/die Vorsitzende: Prof. Dr. A.W. Lohse
Prüfungsausschuss: 2. Gutachter/in: PD Dr. J. Petersen
Prüfungsausschuss: 3. Gutachter/in: Prof. Dr. H. J. Seitz

Table of Contents

Hypothesis and questions ……………………………………………………………... 1

1. Introduction ……………………………………………………………………2

1.1 Autoimmune hepatitis ……………………………………………………………2

1.2 SLA protein and anti-SLA antibodies…………………………………………….. 3

1.2.1 SLA/LP protein and anti-SLA/LP antibodies …………………………………… 3

1.2.2 The structure of SLA/LP protein …………………………………………………3

1.2.3 The function of SLA/LP protein …………………………………………………4

1.2.4 Autoimmunity to SLA/LP protein ……………………………………......... 5

1.2.5 The conservation of SLA/LP protein ……………………………………………5

1.3 The component of a core promoter ………………………………………………..6

1.3.1 Promoter and core promoter …………………………………………………….. 6

1.3.2 The core promoter elements …………………………………………………….. 7

1.3.3 CpG island promoter ……………………………………………………………..10

1.4 Transcription factors ………………………………………………………………. 11

1.4.1 Transcription factor ………………………………………………………………11

1.4.2 The type of transcription factors ………………………………………………… 11

1.4.3 The function of transcription factors ……………………………………………..12

1.4.3.1 The function of general transcription factors ……………………………… 12

1.4.3.2 The function of upstream transcription factors – Sp1, RAP1, Oct-1 ………13

1.4.3.2.1 Sp1 ……………………………………………………………………13

1.4.3.2.2 RAP1 …………………………………………………………………14

1.4.3.2.3 Oct-1 ………………………………………………………………… 15

2. Materials and Methods ……………………………………………………………...17

2.1 Materials ……………………………………………………………………………17

2.1.1 BAC DNA ………………………………………………………………………..17

2.1.2 Cloning vectors ………………………………………………………………….. 17

2.1.3 Cell lines ………………………………………………………………………….17

2.1.4 Enzymes and DNA markers ……………………………………………………...18

2.1.5 Antibodies and proteins …………………………………………………………..18

2.1.6 Kits …………………………………………………………………………………18

2.1.7 Chemicals 18
2.1.8 Some important reagents 19
2.1.9 Common Buffers 19
2.1.10 Primers 23
2.1.11 Laboratory instruments 25
2.1.12 Prediction Programs 25
2.2 Methods 26
2.2.1 Preparation of plasmid DNA 26
2.2.2 Polymerase chain reaction (PCR) 26
2.2.3 DNA electrophoresis 26
2.2.4 DNA extraction and purification 27
2.2.5 Northern Blot 28
2.2.5.1 RNA isolation 28
2.2.5.2 Reverse Transcription (RT) and SLA/LP mRNA-specific PCR (RT-PCR)..28
2.2.5.3 Preparation of Northern Blot probe 28
2.2.5.4 Labeling the Northern Blot probe with isotope 28
2.2.5.5 Northern Blot of Mouse Embryo and mouse tissues 28
2.2.6 Western Blot 29
2.2.6.1 Isolation of protein from different tissues 29
2.2.6.2 Western Blot 29
2.2.7 Cloning and sequencing of mouse SLA/LP promoter fragment 29
2.2.7.1 Stick-end cloning into plasmid vector 29
2.2.7.2 Blunt-end cloning into plasmid vector 30
2.2.8 The generation of 1740bp fragment clone 30
2.2.9 5’-end deletion and 3’-end deletion 32
2.2.10 Point mutation in the sequence of mouse SLA/LP promoter 32
2.2.11 Transient transfection 32
2.2.12 Luciferase assay 33
2.2.13 Gel Shift Assay and Super Gel Shift Assay 34
2.2.13.1 Preparation of double-strands DNA probes 34
2.2.13.2 DNA Binding Reactions 34
2.2.13.3 Super Gel Shift Assay 36
2.2.13.4 Electrophoresis 36
2.2.13.5 Blotting 37

2.2.13.6 Cross-link Transferred DNA to membrane ……………………………….37
2.2.13.7 Detect Biotin-labeled DNA - Chemiluminescent Nucleic Acid Detection Module (Pierce) ………………………………………………..37

3. Results ……………………………………………………………………….. 39
3.1 Western Blot on different mouse tissues ……………………………………. 39
3.2 Northern Blot on Mouse Embryo and different mouse tissues ………………………39
3.3 1740bp SLA/LP fragment and the ability to initiate protein expression ……………. 40
3.4 The typical core promoter elements in the 1740bp SLA/LP fragment ………………41
3.5 5'- or 3'- deletion mutations of 1740bp/pGL3 Basic clone and Luciferase Assay …..41
3.6 Predicted transcription factors and promoter type ……………………………………45
3.7 Point mutations and luciferase assays …………………………………………………..47
3.8 Gel Shift Assays and Super Gel Shift Assays ………………………………………..50

4. Discussion ………………………………………………………………………….53
5. Summary ……………………………………………………………………………..63
6. References …………………………………………………………………………64
7. Abbreviations 78
8. Acknowledgments 80
9. Curriculum Vitae ……………………………………………………………….81
10. Statement ……………………………………………………………………….82

Hypothesis and Questions

Autoimmune hepatitis (AIH) is one of the three putative autoimmune liver diseases that afflict human beings worldwide. The aetiology of AIH is not understood, and it is not clear, which autoantigens drive the pathogenic autoimmune reaction to liver. Thus far, soluble liver antigen/liver pancreas antigen (SLA/LP) is the only known autoantigen, which is specifically recognized only by autoantibodies of AIH patients. Therefore, a role of SLA/LP in the aetiology or pathogenesis of AIH is likely. Autoimmunity to SLA/LP may be driven by aberrant expression of the SLA/LP molecule or linked to the biological function of SLA/LP; however, both the biological function and the regulation of expression of the SLA/LP molecule have not been defined.

To address the regulation of its expression, we studied the transcriptional regulation of SLA/LP gene.

In this study, we mapped the core promoter region of murine SLA/LP gene, and identified several transcription factors, which seem to regulate the expression of the SLA/LP protein. These findings suggest that SLA/LP gene may be a housekeeping gene.

1. Introduction

1.1 Autoimmune hepatitis

Autoimmune hepatitis (AIH) is one of the three putative autoimmune liver diseases afflicting human beings worldwide. The other two are primary biliary cirrhosis (PBC) and primary sclerosing cholangitis (PSC). Neither the etiologies nor immunopathogenetic mechanisms of autoimmune liver diseases have been identified.

Autoimmune hepatitis is a generally progressive, chronic hepatitis of unknown cause. It is more common among women than men, but it occurs globally in children and adults of both sexes in diverse ethnic groups (Pando et al. 1999; Czaja et al. 2002; Yoshizawa et al. 2005). The diagnosis of AIH is based on assessment of clinical and serum biochemical (elevated serum IgG) features, histologic abnormalities and the presence of autoantibodies, such as antinuclear antibody (ANA), smooth muscle antibody (SMA), anti liver-kidney microsome-1 antibody (LKM), perinuclear anti-neutrophilic cytoplasmic antibody (pANCA) and soluble liver antigen/liver pancreas antigen (SLA/LP) (Alvarez et al. 1999). Most AIH patients respond well to immunosuppressive therapy, but have a poor prognosis if untreated (Krawitt 1996). Indeed, the natural history of AIH shows a poor prognosis, with frequent progression to cirrhosis, hepatic insufficiency and, sometimes, to carcinoma in the absence of viral infection. However, the occurrence of carcinoma is rare and formed only in long-standing cirrhosis (Park et al. 2000). Appropriate management can prolong AIH patient survival, improve the quality of life and avoid the need for liver transplantation.

The pathogenesis of AIH is unclear. A conceptual framework postulates an environmental agent that triggers a cascade of T-cell-mediated events directed at liver antigens in a host genetically predisposed to this disease, leading to a progressive necroinflammatory and fibrotic process in the liver (Krawitt 2006). The potential triggers inducing AIH have not been delineated but may include viruses, such as measles virus, hepatitis viruses A, B and C, cytomegalovirus and Epstein-Barr virus (Skoog et al. 2002; Vento et al. 1997; Chiba et al. 2004; Laskus et al 1989; Robertson et al. 1987), or certain drugs (Sterling et al. 1996; Gough et al. 1996; Graziadei et al. 2003). Molecular mimicry of these inducers with autoantigens could play a role in the development of AIH. Thus, the prevalence of autoimmune hepatitis may be higher than reported because of concomitant chronic hepatitis C or B or both (Toda et al. 1997).

The diagnostic hallmark of AIH is the presence of circulating autoantibodies, which target autoantigens that have been characterized to various extents. Autoantibodies are used

as a means of subclassification of AIH into type1 and 2. The main markers of type 1 AIH are antinuclear antibody (ANA) and smooth-muscle antibody (SMA). Type 1 AIH is associated with HLA-DR3 serotype, which is more common among Caucasian patients and in the early-onset, severe form of AIH that often occurs in girls and young women. In HLA-DR3-negative patients, type 1 AIH is often associated with HLA-DR4. HLA-DR4 associated AIH is more common in adults and may be associated with increased incidence of extrahepatic manifestations, milder disease, and a better response to corticosteroid therapy (Krawitt 2006).

Anti liver-kidney microsome-1 antibody (LKM-1) and liver cytosol-1 antibody (LC-1) characterize type 2 AIH. Type 2 AIH is relatively rare and associated with HLA-DRB1 and HLA-DQB1 alleles (Djilali-Saiah et al. 2004), and affects mainly children.

These antibodies can serve as markers of the disease. Although they react with different hepatic proteins that may also be target for tissue-infiltrating effector T lymphocytes (Medina et al. 2003), it is not clear whether autoantibodies have a direct pathogenic role. All the antibodies mentioned above are not specific markers for diagnosing AIH and they can also be found in other diseases.

1.2 SLA/LP protein and anti-SLA/LP antibodies

1.2.1 SLA/LP protein and anti-SLA/LP antibodies

The only antibody, which has been found to be strictly disease-specific in 10 -30 percent of AIH patients, is antibody to soluble liver antigen/liver pancreas antigen (SLA/LP) (Wies et al. 2000; Baeres et al. 2002). In overlap syndrome of AIH with primary biliary cirrhosis (PBC) or primary sclerosing cholangitis (PSC), anti-SLA/LP can also be positive.

SLA/LP is a cytosolic protein of about 50 kDa expressed in enzymatically active organs such as liver, pancreas, kidney, lung, testis and overexpressed in activated lymphocytes (Wies et al. 2000).

1.2.2 The structure of SLA/LP protein

A structure model of SLA/LP antigen was published recently (Kernebeck et al. 2001). The sequence of SLA/LP is compatible with an architecture of the superfamily of pyridoxal phosphate (PLP, vitamin B6)-dependent transferases. It was identified as a gene encoding 474 amino acid residues. The main antigenic region crucial for recognition of anti-SLA/LP lies between amino acids 371 and 409. This region shows substantial homologies with various microbial antigens, including proteins of Rickettsia species, human herpesvirus 6,

and cytomegalovirus (Wies et al. 2000). However, the homologous microbial sequences are only poorly recognised by SLA/LP autoantibodies (Herkel et al. 2002). Nevertheless, molecular mimicry of homologous proteins from other species may be a trigger of SLA/LP autoimmunity.

The structure of SLA/LP is shown in Fig 1.

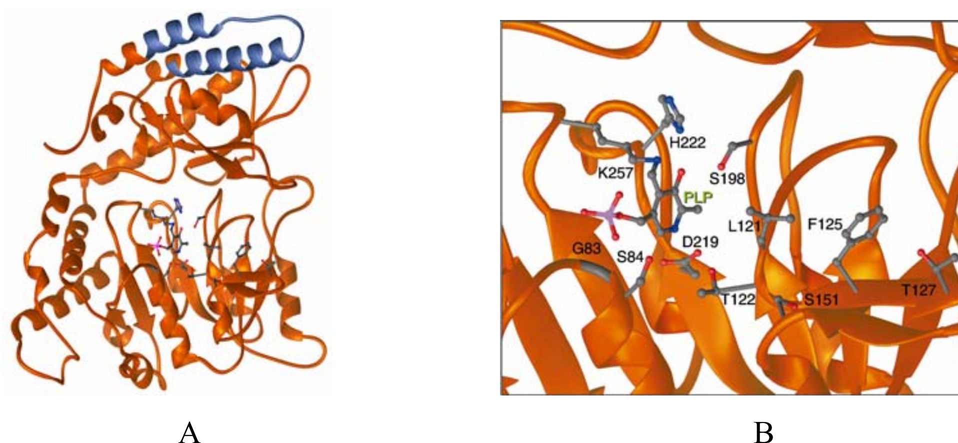

Fig 1. The SLA/LP protein structure (Kernebeck et al. 2001).

A. Three-dimensional model of SLA/LP. A ribbon representation showing the backbone of the 3-dimensional model of the soluble liver antigen. The dominant epitope region is colored in grey

B. Close up of the active site of the 3-dimensional model of SLA/LP. Residues known to bind PLP15 and that are conserved in SLA/LP are depicted (G83, S84, S198, D219, H222, and K257, which is covalently bound to PLP). The conserved residues L121, T122, F125, T127, and S151 belong to a region that is known to act as a channel between the active site and the solvent.

1.2.3 The function of SLA/LP protein

The primary biological function of SLA/LP remains unclear. Because the SLA/LP molecule was found to be associated with the UGA tRNP(Ser)Sec complex (Gelpi et al. 1992; Costa et al. 2000), which facilitates the co-translational incorporation of selenocysteine into proteins, it has been speculated that the SLA/LP molecule may have a role in selenoprotein metabolism; the specialized UGA tRNA is initially charged with serine to form seryl-tRNA, which then is enzymatically converted to selenocysteyl-tRNA^{Sec}. However, there is no direct experimental evidence for such a role of the SLA/LP molecule so far. Nevertheless, a fold recognition study predicted the SLA/LP tertiary structure by comparison to known protein structures to be that of a pyridoxal phosphate (PLP)-dependent transferase (Kernebeck et al. 2001), which is compatible with a role in selenoprotein metabolism. The active site was proposed to be a cavity with a channel, formed by dimerisation of two SLA/LP molecules (Kernebeck et al. 2001; Scarsdale et al. 1999). In the three dimensional model, five amino

acids of monomer A (L88, T89, F92, T94, and S118) as well as 2 amino acids of monomer B (P251 and G252) are involved in dimerisation and the amino acids critical for binding and orientation of the co-enzyme PLP were identified to be G50, S51, S165, D186, H189, K224 (Kernebeck et al. 2001) (residue numbering according to GenBank accession number NP_722547). The latest data shows that SLA/LP, together with SECp43, formed a complex with selenocysteine (Sec) $tRNA^{(Ser)Sec}$, and regulate selenoprotein expression and firmly linked these proteins to the pathway of selenoprotein biosynthesis (Xu et al. 2005).

The role of SLA/LP in AIH also remains speculative. It may be involved in the pathophysiology of autoimmune hepatitis. Alternatively, its substrate or metabolite may be related to the pathogenesis, or it may play a role in immune regulation.

1.2.4 Autoimmunity to SLA/LP protein

The SLA/LP molecule is a cytoplasmic protein and it is not clear how SLA/LP autoantibodies may recognise an intracellular protein; a possible translocation of SLA/LP molecules to the cell surface has not been examined. Alternatively, liver cell damage may be mediated by SLA/LP-specific T lymphocytes; although specific T cells have not been described yet, a pathogenic role for specific T cells is likely, given the highly selected phenotype of SLA/LP autoantibodies. Be that as it may, it is also possible that autoimmunity to SLA/LP is only an epiphenomenon of liver cell damage, and not involved in pathogenesis. However, preliminary findings suggest that, at least in mice, hepatic inflammation and liver cell damage can be induced by autoimmunisation to SLA/LP (Herkel, personal communication). It is also possible that the substrate or the metabolite of the SLA/LP enzyme may be a crucial antigen, which triggers the autoimmunity to SLA/LP protein (Kernebeck et al. 2001).

1.2.5 The conservation of SLA/LP protein

The SLA/LP molecule was highly conserved in evolution, and sequences from various species, including man, mouse, zebrafish, fruit fly and worm, display high degrees of similarity or homology, suggesting an indispensable function of the molecule. The highest degree of similarity was found between the human and the mouse amino acid sequences; both species also have a highly similar exon/intron structure in the SLA/LP gene. Moreover, both mammalian species seem to generate similar variant proteins by differential splicing of exon 2 (Wang et al. 2006).

The human SLA/LP gene sequence of approximately 39kb, which maps to chromosome 4p15.2, is organised in 11 exons, of which 10 or 11 are translated, depending on the splice variant (The accession number is NM-016955 and NM-153825.). The mouse SLA/LP gene sequence, which maps to mouse chromosome 5qC1, spans 28.5 kb and the mouse SLA/LP gene is organized into 11 exons like the human homologue (NM-172490). Homologous molecules were identified in several biological model organisms, which showed a high degree of similarity, notably at those residues that are of functional importance. The only domain that lacks significant homology is the major antigenic epitope of the human protein sequence recognised by autoantibodies from AIH patients. Thus, it appears that SLA/LP autoimmunity is specific for the self-antigen and not for homologous sequences from other eukaryotic species. The possibility that SLA/LP autoimmunity might be driven by homologous proteins from parasites is hence quite unlikely.

SLA/LP-homologous proteins are only found in eukaryotes and archaebacteria, but not in eubacteria (Herkel et al. 2002). Nevertheless, a few homologous sequences from bacterial or viral proteins with some degree of similarity to the antigenic epitope of the SLA/LP protein do exist; however, these are not recognised by SLA/LP autoantibodies (Herkel et al. 2002). Likewise, the corresponding sequence of an archaebacterial SLA/LP-homologue is also not recognised by SLA/LP autoantibodies (Herkel et al. 2002). Therefore, SLA/LP autoimmunity in patients is very likely driven by the self-SLA/LP molecule rather than by a mechanism that involves molecular mimicry.

1.3 The component of a core promoter

Accurate prediction of promoters is fundamental understanding gene expression patterns, cell specificity and development. Promoter can function not only to bind RNA polymerase, but also to specify the places and times that transcription can occur from that gene. In eukaryotes, promoters are recognized by specific transcription factors.

1.3.1 Promoter and core promoter

A promoter is a DNA sequence that enables a gene to be transcribed. The promoter is recognized by RNA polymerase, which then initiates transcription. In RNA synthesis, promoters are a means to demarcate which genes should be used for messenger RNA creation and, by extension, control which proteins the cell manufactures. Promoters represent critical elements that can work in concert with other regulatory regions (enhancers, silencers, boundary elements/insulators) to direct the level of transcription of a given gene.

A promoter is composed of three parts. The first part is the core promoter, which is the minimal stretch of contiguous DNA sequence that is sufficient to direct accurate initiation of transcription by the RNA polymerase II machinery (Butler et al. 2002). The second part is the proximal promoter ranging 200-300bp immediately upstream of the core promoter, which contains multiple transcription factor binding sites, responsible for transcription regulation. The third part is the distal part of the promoter known as enhancer/silencer element, which is located further upstream and may also include transcription factor binding sites (Lemon et al. 2000; Smale 2001). The generic structure of a typical promoter is shown in Fig 2.

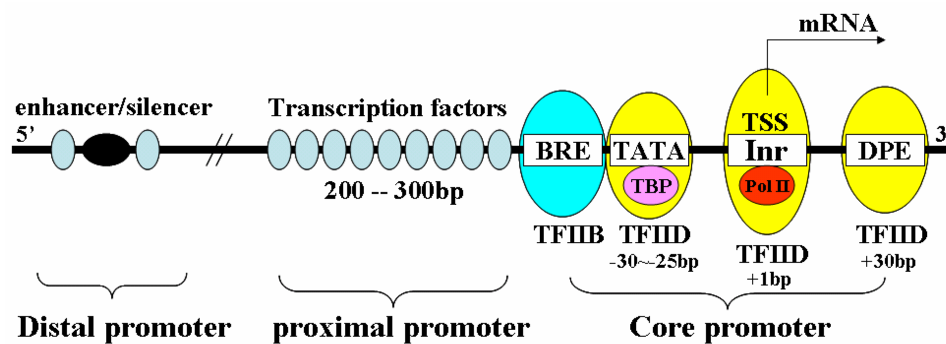

Fig 2. The generic structure of a typical promoter. This diagram shows the core promoter, proximal promoter and distal promoter and the main elements inside.

1.3.2 The core promoter elements

Typically, the core promoter encompasses the site of transcription initiation and extends either upstream or downstream for an additional ~35 nucleotides. Thus, in many instances, the core promoter will comprise only about 40 nucleotides. There are several sequence motifs—which include the TATA box, initiator (Inr), TFIIB recognition element (BRE), and downstream promoter element (DPE)—that are most commonly found in core promoters (Smale et al. 2003). There are some newly found core promoter elements, such as downstream core element (DCE) discovered in the human β-globin promoter (Lewis et al. 2000), motif ten element (MTE) conserved from Drosophila to humans (Lim et al. 2004). It is important to note that each of these core promoter elements is found in some but not all core promoters. For instance, TATA-containing core promoter is about 32% of 1031 potential promoter regions in humans (Suzuki et al. 2001), 10 to 20% of all known human promoters (Gershenzon et al. 2005), and 43% of 205 core promoters in Drosophila (Kutach et al. 2000). In some large groups of genes, like housekeeping genes, oncogenes and growth factor genes, TATA box is often absent, and the corresponding promoters are referred to as TATA-less promoters.

TATA box (Goldberg-Hogness box) was the first eukaryotic core promoter motif to be identified (Goldberg 1979; Breathnach et al. 1981). In metazoans, the TATA box is typically located about 25-30 nucleotides upstream of the transcription start site (TSS) (Butler et al. 2002). TATA box-binding protein (TBP) is the predominant binding protein of TATA box. During the formation of the active eukaryotic initiation complex, transcription factor IID (TFIID) binds to the TATA box through its TBP subunit, then RNA polymerase II bind to TATA box through TFIID with the help of TFIIB and other factors. The RNA polymerase II is now competent to transcribe mRNA from the gene.

The initiator (Inr) element encompasses the transcription start site (TSS), from -3 to +5 with the consensus sequence as PyPyA^{+1}NT/APyPy (Smale et al. 1990; Javahery et al. 1994). The A^{+1} position is designated at the +1 start site because transcription commonly initiates at this nucleotide. Only a subset of the pyrimidines at the -2, +4, and +5 positions appears to be essential for Inr activity, but the activity increases with increasing numbers of pyrimidines in these positions (Javahery et al. 1994; Lo et al. 1996). Based on a database analysis, the consensus sequence, PyCA^{+1}NTPyPy, is more common in mammals (Bucher 1990; Corden et al. 1980), while TCA^{+1}G/TTPy is more common in Drosophila (Arkhipova 1995; Kutach et al. 2000; Ohler et al. 2002).

Transcription of genes with promoters containing a TATA box or initiator element normally begins at a well-defined initiation site. However, transcription does not need to begin at the +1 nucleotide for the Inr to function. Transcription initiates, more generally, at a single site or in a cluster of multiple sites in the vicinity of the Inr. RNA polymerase II has been redirected to alternative start sites by reducing ATP concentration within a nuclear extract, by altering the spacing between the TATA and Inr in a promoter containing both elements, and by dinucleotide initiation strategies (O'Shea-Greenfield et al. 1992; Kadonaga 1990; Zenzie-Gregory et al. 1992). In all of these studies, the Inr continued to increase the efficiency of transcription initiation from the alternative sites.

Transcription of genes with a promoter not containing a TATA box or an initiator has been shown to begin at any one of multiple possible sites over an extended region, often 20 to 200 base pairs in length. As a result, such genes give rise to mRNAs with multiple alternative 5' ends. These genes, which generally are transcribed at low rates, e.g. genes encoding the enzymes of intermediary metabolism, are often refered to as "housekeeping genes". Most genes of this type contain a CG-rich stretch of 20 to 50 nucleotides within $\approx$ 100 base pairs upstream of the start-site region. A transcription factor called Sp1 recognizes these CG-rich sequences (Lodish et al. 2000).

Inr elements are found in both TATA-containing promoters with 61.9% of percentage as well as TATA-less promoters with 45.4% of percentage (Gershenzon et al. 2005). A variety of factors have been found to interact with the Inr element. For instance, several studies have confirmed that TFIID specifically interacts with the Inr (Wang et al. 1993; Verrijzer et al. 1995; Bellorini et al. 1996; Burke et al. 1996). In vitro experiments ascribed two distinct activities to an Inr (Smale et al. 1989): (i) the ability to independently direct RNA polymerase II to initiate transcription from a specific, internal position; and (ii) the ability to be activated in the absence of TATA by an upstream activator element, resulting in high levels of accurate transcription.

The downstream promoter element (DPE) was mainly studied in Drosophila (Kutach et al. 2000). The DPE was identified as a downstream core promoter motif that is required for the binding of purified TFIID to a subset of TATA-less promoters (Burke et al. 1996; Kadonaga 2002; Butler et al. 2002). It was shown that DPE is conserved from Drosophila to human (Burke et al. 1997). The consensus sequence of DPE is A/G G A/T C G T G (Burke et al. 1996). The DPE acts in conjunction with the Inr, and the core sequence of the DPE is located at precisely +28 to +32 relative to the A+1 nucleotide in the Inr motif (Kutach et al. 2000). TFIID binds cooperatively to the Inr and DPE motifs, as mutation of either the Inr or the DPE results in loss of TFIID binding to the core promoter (Burke et al. 1996). The DPE is found most commonly in TATA-less core promoters. With naturally occurring TATA-less core promoters, mutation of the DPE motif results in a 10- to 50-fold reduction in basal transcription activity (Burke et al. 1996, 1997; Kutach et al. 2000).

The TFIIB recognition element (BRE) is a TFIIB binding site that locates immediately upstream of some TATA boxes in some TATA-containing promoters (Lagrange et al. 1998), and can increase the affinity of TFIIB for the core promoter.

The core promoter elements are shown in Fig 3.

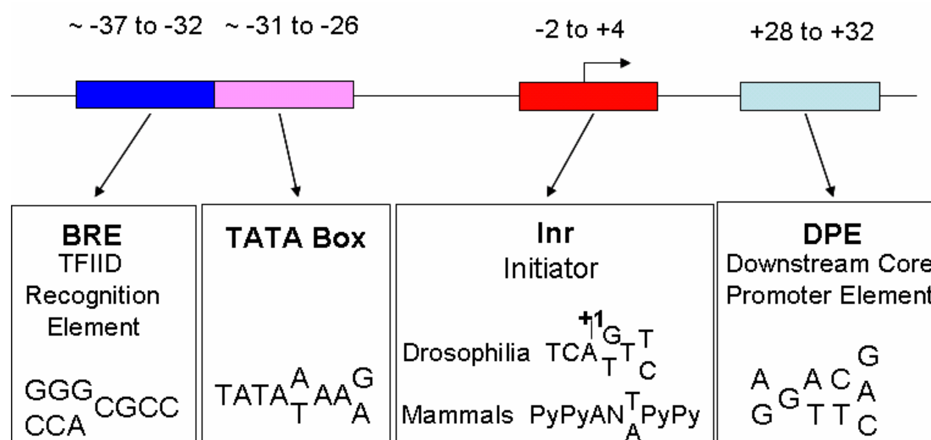

Fig 3. Core promoter motifs (Smale et al. 2003). This diagram depicts some of the sequence elements that can contribute to basal transcription from a core promoter. Each of these sequence motifs is found in only a subset of core promoters. A particular core promoter many contain some, all, or none of these elements. The TATA box can function in the absence of BRE, Inr, and DPE motifs. In contrast, the DPE motif requires the presence of an Inr. The BRE is located immediately upstream of a subset of TATA box motifs. The DPE consensus was determined with *Drosophilia* core promoter. The Inr consensus is shown for both mammals and *Drosophilia*. (Py= pyrimidine)

1.3.3 CpG island promoter

Whether a promoter is located in CpG islands or not is also very important for transcriptional regulation. CpG islands are defined as dispersed regions of DNA with high frequency of CpG dinucleotide relative to the bulk genome (Gardiner-Garden et al. 1987; Larsen et al. 1992). When CpG islands remain unmethylated, TF-binding sites can be recognized by TF. In contrast, when methylated, the presence of 5-methylcytosine in CpG islands interferes with the binding of TFs and thus suppresses transcription. CpG islands are often located around the promoters of housekeeping genes, growth factor genes, oncogenes and other frequently expressed genes in cells (Larsen et al. 1992; Cross et al. 1995). CpG islands, which generally range in size from 0.5 to 2 kbp, contain promoters for a wide variety of genes. It has been estimated that, in mammals, CpG islands are associated with approximately half of the promoters for protein-coding genes (Suzuki et al. 2001; Antequera et al. 1993). Despite the prevalence of promoters associated with CpG islands, the elements that are responsible for their core promoter function remain poorly defined. A widely held opinion is that CpG islands usually lack consensus or near-consensus TATA box, DPE elements, or Inr core promoter elements (Blake et al. 1990). From the core promoter perspective, CpG islands may contain multiple weak core promoters rather than a single strong promoter (Butler et al. 2002). As a consequence, they are often characterized by the presence of multiple weak transcription start sites that span a region of 100bp or more. The transcription start sites can coincide with sequences exhibiting weak homology to the Inr consensus or can be unrelated to this sequence. Mutation in the vicinity of the start site can lead to the use of alternative start sites, but promoter strength is often unaffected. In general, it has been difficult to identify core promoter elements within CpG islands that are essential for promoter function (Smale et al. 2003). One common feature of CpG islands is the presence of multiple GC box motifs that are bound by transcription factor Sp1 and related transcription factors (Brandeis et al. 1994; Macleod et al. 1994; Blake et al. 1990). The presence of Sp1 binding sites in CpG islands is particularly notable. It has been found that

Sp1 binding sites in conjunction with an Inr motif can activate transcription in the absence of a TATA box (Smale 1990; Emami et al. 1995). Hence, it is possible that CpG islands promoters consist of multiple Sp1+Inr pairs that collectively generate the array of start sites. However, of the promoters with CpG islands, still about 6.9% have a TATA box, about 45.2% have Inr, about 24.3% have DPE, and about 33.4% have BRE elements (Gershenzon et al. 2005).

Statistical sequence analysis of 8793 human promoters revealed that: (1) the majority of promoters (74.3%) have at least one of four core promoter elements at their functional position and 44.1% have only one element. The portion of the TATA-containing promoter is just from 10 to 20% of all known human promoters. (2) One-fourth of all promoters do not have any of the four core promoter elements suggesting the existence of other yet undiscovered core elements. (3) The statistical significances of the occurrence frequency of the DPE and BRE elements at their experimentally defined functional positions are high, indicating that considerable amount of human genes use these elements for the transcription. (4) The high percentage and statistical significance of BRE, especially in CpG-containing and multiple-transcription start site (MSS) promoters, suggests that this element may be functional in many promoters including TATA-less promoters (Gershenzon et al. 2005). An analysis of the potential promoter regions of 1031 kinds of human genes showed that the core promoter elements appears with TATA box at 32%, Inr at 85%, GC box at 97% and 48% of the promoters were located in CpG islands (Suzuki et al. 2001).

1.4 Transcription factors

1.4.1 Transcription factor

In molecular biology, a transcription factor is a protein that binds DNA at a specific promoter or enhancer region or site, where it regulates transcription. As a component of promoter, transcription factors (TF) usually locate in proximal promoter part and distal promoter part. Transcription factors can be selectively activated or deactivated by other proteins, often as the final step in signal transduction.

1.4.2 The type of transcription factors

There are three classes of transcription factors. The first is general transcription factors, which are involved in the formation of a preinitiation complex. The most common general transcription factors are abbreviated as TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH. They

are ubiquitous and interact with the core promoter region surrounding the transcription start site(s). The second type is upstream transcription factors, which are unregulated proteins that bind to a cis-regulatory element (such as an enhancer or repressor sequence) somewhere upstream of the initiation site to stimulate or repress transcription either directly or indirectly. The third type is inducible transcription factors, which are similar to upstream transcription factors but require activation or inhibition.

The transcription factors in higher eukaryotes can also be divided into three general groups according to their binding expression: (1) General transcription factors (TFIID, TFIIB, TFIIA, TFIIH, TFIIE, and TFIIF) which together with RNA polymerase II form a basal transcription complex. These factors are expressed in all cell types. (2) Transcription factors that bind to specific DNA sequences and express in all cell types. These include Sp1, CCAAT-box binding protein, RAP1 and many others. These ubiquitously expressed transcription factors mostly activate transcription of many genes including tissue specific genes. (3) Tissue specific transcription factors also bind to specific DNA sequences but express in specific cell types, bind to tissue-specific enhancer element and activate transcription of tissue specific genes.

1.4.3 The function of transcription factors

1.4.3.1 The function of general transcription factors

Transcription requires the interaction of RNA polymerase with promoter DNA. In eukaryotic cells, there are three different types of RNA polymerases, each having particular functions and properties (Valenzuela et al. 1976). RNA polymerase I is responsible for transcribing the large ribosomal RNAs; RNA polymerase II transcribes messenger RNA precursors; and RNA polymerase III transcribes small RNAs such as transfer RNAs, 5S ribosomal RNA and other small sequences. However, none of the eukaryotic RNA polymerase can bind efficiently to DNA. Hence, transcription factors, the families of DNA binding proteins, first bind to DNA and interact with the RNA polymerase to initiate RNA synthesis.

A fundamental step of the transcription initiation is an interaction of the basal transcription machinery [also named pre-initiation complex (PIC)]. RNA polymerase II is a multisubunit enzyme that catalyzes the synthesis of mRNA from the DNA template. Accurate and efficient transcription from the DNA template (core promoter) requires the polymerase along with auxiliary factors, such as basal or general transcription factors. Of the

general transcription factors, TFIID always plays the central role in successful transcription (Burley et al. 1996; Burke et al. 1997), acting in cooperation with the core promoter elements and/or specific TFs (Nikolov et al. 1997; Hampsey 1998; Lemon et al. 2000). The TFIID consists of TATA Binding Protein (TBP) subunit and at least 12 transcription associated factors (TAFs) (Green 2000). In the TATA box-containing promoters, TBP binding starts the process of the pre-initiation complex (PIC) formation. In the TATA box-less promoters, TAFs bind to DNA and /or other TFs in order to involve TFIID (and TBP) in PIC (Burke et al. 1997; Zenzie-Gregory 1993; Martinez et al. 1995; Tsai et al. 2000).

Thus, the function of general transcription factors is to form the basal transcription machinery and initiate transcription.

1.4.3.2 The function of upstream transcription factors – Sp1, RAP1, Oct-1

The upstream transcription factors and inducible transcription factors, such as Sp1, RAP1 and Oct-1, bind to specific cis-elements, regulate the transcription process in the way of either activation or repression.

1.4.3.2.1 Sp1

Specificity protein 1 or stimulating protein 1 (Sp1), the first transcription factor identified, was isolated from HeLa cells and was originally cloned as a factor that binds to the SV40 early promoter (Dynan et al. 1983; Gidoni et al. 1984). It is the founding member of a growing Sp family which contain a highly conserved DNA-binding domain composed of three conserved Cys2His2 zinc fingers close the C-terminus and serine/threonine- and glutamine-rich domains in their N-terminal regions (Briggs et al. 1986; Suske 1999; Bouwman et al. 2002; Kaczynski et al. 2003; Li et al. 2004). The glutamine and serine/threonine rich N-terminus of Sp1 strongly activates transcription, while the C-terminus containing three zinc fingers activates transcription poorly by themselves, but is essential for synergistic activation of transcription (Kadonaga et al.1987; Courey et al. 1989; Emami et al.1995). The three zinc fingers bind GC or GT boxes in the promoter or enhancer region of many genes (Kadonaga et al. 1987). Within Sp family, Sp1 has the ability to form multimers (Yu et al. 2003), and typically functions as an activator of transcription (Suske 1999; Bouwman et al. 2002; Li et al. 2004). Like many activators, Sp1 requires the transcription factor IID (TFIID) complex for efficient stimulation of transcription in vitro (Smale et al. 1990). Sp1 is ubiquitously expressed in mammalian cells and participates in regulating the expression of genes involved in almost all cellular processes (Cawley et al. 2004), such as

cell cycle regulation (Karlseder et al. 1996; Black et al. 1999; Kavurma et al. 2003), chromatin remodelling (Jongstra et al. 1984; Ellis et al. 1996), prevention of CpG island methylation (Brandeis et al. 1994; Macleod et al. 1994), and apoptosis (Li-Weber et al. 1998; McClure et al. 1999; Kavurma et al. 2001,2003).

It has been proposed that many TATA-less and GC-rich promoters bind one or more Sp1 molecules to recruit specific cofactors such as TATA-binding protein associated factors (TAFs), which subsequently interact with TF IID to initiate transcription (Pugh et al. 1990; Goodrich et al. 1994), as shown in Fig 4.

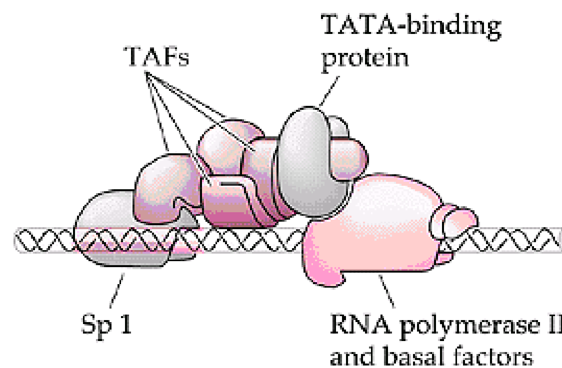

Fig 4. Possible configuration for TF mediating RNA polymerase II binding to a TATA-less promoter containing an Sp1-binding site (Pugh and Tjian 1991; Comai et al.1992).

1.4.3.2.2 RAP1

The repressor activator protein 1 (RAP1) is a multifunctional, sequence-specific, DNA-binding protein involved in diverse cellular processes such as transcriptional activation and silencing, translation, nutrient transport, glycolysis, mating type regulation (Capieaux et al. 1989), and is an essential factor for telomere length regulation and maintenance. Furthermore, its activity with respect to certain genes is regulated by growth conditions (Henry et al. 1990).

RAP1 is mainly studied in yeast. A high affinity RAP1 consensus binding site of the form 5'(A/G)(A/C)ACCCANNCA(T/C)(T/C)3', where N is any nucleotide, was proposed (Buchman et al. 1988). However, it is clear that not all strong RAP1 binding sites are perfect matches to this consensus (Shore et al. 1987; Capieaux et al. 1989; Chambers et al. 1989; Devlin et al. 1991; Fantino et al. 1992). The positions 2 to 7 form the core of the RAP1 binding site and positions 4 and 5 are absolutely critical for RAP1 binding (Graham et al. 1994). The RAP1 DNA-binding sequence: 5' ACACCCATACATTT 3' is called upstream activator sequence [(UAS)rpg], while 5' ACACCCACACACCC 3' is called telomere consensus sequence (Idrissi et al. 1998). These two sequences differ in their activation potential. When assayed as direct repeats, the UASrpg showed a strong synergistic effect,

which was orientation-dependent. In contrast, the telomeric sequence showed a much lower synergism, with no dependence on orientation (Idrissi et al. 1999). This was confirmed by telomeric RAP1-binding sequence 5' GGTGTGTGGGTGT 3' (Konig et al. 1997) which is the same as the anti-parallel (reverse and complement) sequence of the telomeric sequence.

RAP1 binds to DNA through two Myb-type helix-turn-helix motifs (Konig et al. 1996). The amino acid sequence of RAP1 in *Saccharomyces cerevisiae*, TAZ1 (transcriptional adaptor zinc-binding domain) in *Schizosaccharomyces pombe*, and human TRF1 (telomeric repeat binding factor 1) and TRF2 show similarities to each other and to the DNA-binding motif in the c-Myb family of transcription factors (Konig et al. 1997). C-Myb proteins typically consist of three tandem repeats of the Myb DNA-binding motifs, where at least two, which are tandemly repeated GGTGT, are required for sequence-specific DNA recognition (Tanikawa et al. 1993; Wahlin et al. 2000). A human RAP1 homolog was shown to localize to chromosome ends, bind telomeric DNA through TRF2 and be involved in telomere length regulation, but its function in transcriptional regulation is not known (Li et al. 2000).

1.4.3.2.3 Oct-1

Octamer transcription factor-1 (Oct-1) is a member of the POU transcription factor family (Verrijzer 1993). The POU domain is the DNA binding domain of a class of transcription factors involved in developmental regulation. It was initially discovered as a conserved region in three mammalian transcription factors, **P**it-1, **O**ct-1/Oct-2 and **U**nc-86 (Herr et al. 1988). POU domain is characterized by the presence of a bipartite DNA-binding domain (POU domain). The POU domain contains a POU-specific domain and a POU homeodomain (Herr et al. 1988; Sturm et al. 1988). Both these subdomains have a helix-turn-helix motif, acting not only as a DNA-binding domain but also as a protein-protein interaction domain. The DNA binding specificity is contributed by both components of the POU domain (Brugnera et al. 1992). Members of the POU transcription factor family are involved in a broad range of biological processes. Several members of the POU gene family have been demonstrated to exert critical functions in the regulation of cell-type-specific gene expression, DNA replication, cellular proliferation, hormonal signals pathway, determination of cell identity and developmental control (Rosenfeld 1991; Ruvkun et al. 1991; Scholer 1991; Chandran et al. 1999).

Oct-1 is known as a ubiquitous nuclear protein expressed in a variety of tissues and cell types (Sturm et al. 1987). It activates the octamer motif (5'-ATGCAAAT-3') containing gene promoters that are ubiquitously as well as tissue-specifically expressed genes such as histone

H2B (Fletcher et al. 1987), the small nuclear RNA gene (Murphy et al. 1989), and Ig heavy chain and kappa light chain genes (Franke et al. 1994). A number of transcription factors have been identified to interact with the POU domains of Oct-1 such as TBP, TFIIB, HMG2, and Oct-binding factor-1 (OBF-1) also referred to as Oct-1-associated coactivator (OCA-B) (Zwilling et al. 1994, 1995; Nakshatri et al. 1995; Gstaiger et al. 1996; Strubin et al. 1995; Luo et al. 1995). It is believed that Oct-1 may participate in tissue-specific gene expression by interaction with either other transcription factors (Voss et al. 1991; Kutoh et al. 1992) or tissue-specific coactivators (Luo et al. 1992; Strubin et al. 1995). Isoforms of Oct-1 also contribute to tissue-specific expression (Pankratova et al. 2001; Zhao et al. 2004).

Oct-1 plays multiple roles in cells in many fields. It may act as either a positive or a negative regulator of gene transcription and DNA replication (Verrijzer et al. 1993; Ryan et al. 1997). The ability of Oct-1 to regulate expression of proteins involved in cell cycle regulation (Brockman et al. 2005; Fletcher et al. 1987; Magne et al. 2003), apoptosis (Hirose et al. 2003; Jin et al. 2001), immunity (Cron et al. 2001; Franke et al. 1994; Iademarco et al. 1992, 1993; Osborne et al. 2001, 2004; Prabhu et al. 1996; Zhang et al. 1999) and Oct-1 activation in response to stress signals (Hirose et al. 2003; Jin et al. 2001; Schild-Poulter et al. 2003; Zhao et al. 2000), including viral infection (Advani et al. 2003), suggests an important role for Oct-1 in a defense mechanism against cellular stress (Mesplede et al. 2005). Oct-1 also involves in the regulation of some housekeeping gene functions (Witt et al. 1997; Ryan et al. 1997). It is interesting that Oct-1 is induced after cells are exposed to multiple DNA damaging agents and therapeutic agents, not only in the increased protein level but also in the activity of Oct-1 DNA binding to its specific consensus sequence. This indicates that Oct-1 might participate in cellular response to DNA damage, particularly in p53-independent gene activation (Zhao et al. 2000).

2. Materials and Methods

2.1 Materials

2.1.1 BAC DNA

The mouse SLA BAC DNA clone RP23-70D14 was used for generating SLA promoter fragment.

2.1.2 Cloning vectors

pBlueScript SK（+）vector : Stratagene，#212205

pET-30a (+) vector, Novagen, #69909-3

pGL3-Basic vector, Promega, #E1751, Mannheim

pSV-β-Galactosidase control vector, Promega, #E1081, Mannheim

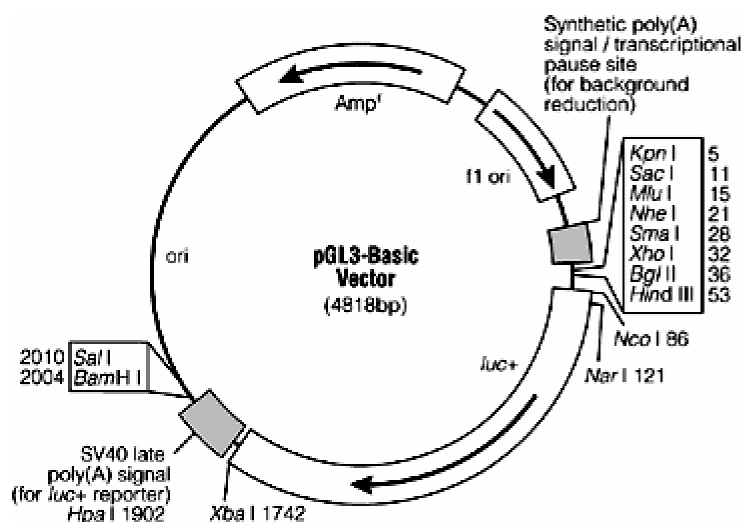

Fig 5. The luciferase reportor vector, pGL3 Basic vector, which lacks of promoter, provides a basis for the quantitative analysis of factors that potentially regulate mammalian gene expression. These factors may be cis-acting, such as promoters and enhancers, or trans-acting, such as various DNA-binding factors (Promega).

2.1.3 Cell lines

The following three cell lines were used for luciferase assay:

HEK293: human embryonic kidney epithelial cell line

Hepa1-6: mouse hepatocellular cell line

RAW264.7: murine macrophage cell line

2.1.4 Enzymes, DNA markers

Kpn I, XmaI, EcoR I, Sau I, Spe I, Xho I, BamH I, EcoR I, Pst I, T4 DNA ligase, Antarctic Phosphatase were bought from New England Biolabs (Frankfurt am Main)

Klenow enzyme, DNA Molecular Weight Marker IV, VII, and VIII were bought from Roche (Mannheim).

2.1.5 Antibodies and proteins

Anti-Sp1: rabbit anti-human polyclonal IgG, EMSA tested, Upstate Cell Signaling Solution, #07-645

Anti-RAP1: Clone 4C8/1, mouse anti-human monoclonal IgG_{2b}, not EMSA tested, Upstate Cell Signaling Solution, #05-911

Anti-Oct-1: Clone YL15, mouse ascites, Upstate Cell Signaling Solution, #05-240,

Recombinant human Sp1 protein (rhSp1): Promega, #E6391

HeLaScribe Nuclear Extract, Gel Shift Assay Grade, Promega, #E3521

2.1.6 Kits

ß-Galactosidase Enzyme Assay System with Reporter Lysis Buffer, Promega, #E2000
DNA Ligation Kit Ver2.1, TaKaRa, #6022
Expand High Fidelity PCR System, Roche, #1732641
Gel Shift Assay System, Promega, #E3300
HiSpeed Plasmid Maxi Kit, Qiagen, #12663
LightShift Chemiluminescent EMSA Kit, PIERCE, #20148
Luciferase Assay System, Promega, #E1500
Mouse Embryo MTN Blot, BD Biosciences Clontech, #636810
Mouse Multiple Tissue Northern Blot, BD Biosciences Clontech, #636808
Prime-It II Random Primer Labeling Kit, Stratagene, #300385
ProSTAR First-Strand RT-PCR Kit, Stratagene, #200420
QIAquick Gel Extraction Kit(250), Qiagen, #28706
QIAprep Spin Miniprep Kit(50), Qiagen, #27104
Quick Ligation Kit, New England BioLabs, #M2200S
REDTaq Readymix PCR Reaction Mix, Sigma, #R2523

2.1.7 Chemicals

All chemicals were delivered from one of these following companies:

Merck Eurolab GmbH, Frankfurt; Carl Roth GmbH, Karsruhe; Sigma-Aldrich Chemie GmbH, Taufkirchen; Roche Diagnostics GmbH, Mannheim; J.T.Baker from Th. Geyer Hamburg GmbH & Co. KG.

2.1.8 Some important reagents

Acrylamide, J.T.Baker, #4081-00
Ammonium peroxodisulphate (APS): Roth, #9592.2
Bio-Rad Protein Assay Dye Reagent Concentrate, Bio-Rad, #500-0006
Chloroform/Isoamyl alcohol (24:1), Serva, #39554
Complete Mini Protease inhibitor cocktail tablets, Roche Diagnostics, #11836153001
DAB Substrate, Roche, #1718096
DMEM: Gibco, #31966-21
Deoxynucleotide Mix, Sigma, #D-7295
Dimethylsulfoxid for molecular biology(DMSO), ROTH, #A994.2
Fetal Bovine Serum, Biochrom KG, #S0115,
FuGENE6 Transfection Reagent, Roche, #11814443001(1ml)
Hybond-N+ membrane, Amersham, #RPN203B
α-^{32}P-dCTP, Amersham Biosciences
Phenol/Chloroform/Isoamyl alcohol (25:24:1), Roth, #A156.1
Poly(dI-dC).(dI-dC), Sigma, #P4929-5UN
ProbeQuant G-50 micro columns, Amersham Biosciences, #275335-01
Protein88, Novartis Nutrition GmbH, #2720796
Rotiphorese Gel 30 (37.5:1): Roth, # 3029.1
RPMI 1640 medium: Gibco, #61870-010
TEMED: Roth, #2367.1
TRIReagent: Sigma, #T9424-100ml
Trypsin-EDTA: Gibco, #25300-054
XL2-Blue Ultracompetent cells, Stratagene, #200150
XL10-Gold Ultracompetent cells, Stratagene, #200314

2.1.9 Common Buffers

Low TE buffer

10mM Tris/0.1mM EDTA (pH8.0)

1% Agarose gel – 150ml

Agarose	1.5g
0.5×TBE buffer	150ml
10mg/ml Ethidiumbromid	3µl

10×DNA Loading Buffer

30% Ficoll 400

100mM Ethylendiaminetetraacetic acid (EDTA, pH8.0)

1% Sodium Dodecyl Sulfat (SDS)

0.25% Bromphenolblue

0.25% Xylene Cyanole FF

in H_2O

1×LB medium (pH7.0) – 1000ml

Tryptone	10g
Yeast extract	5g
NaCl	10g

Distilled H_2O to 1000ml.

LB Agar – 500ml

Tryptone	5g
Yeast extract	2.5g
NaCl	5g
Agar	10g
Distilled H_2O to	500ml.

1×PBS Dulbecco's—1000ml		final concentration
NaCl	8.0g	137mM
Na_2HPO_4	1.15g	8.1mM
KCl	0.2g	2.7mM
KH_2PO_4	0.2g	1.47mM

Distilled H_2O to 1000ml.

10×TBE buffer – 1000ml

Tris	108g
Boric Acid	55g
EDTA-Na2	7.44g
Distilled H_2O to	1000ml.

0.5×TBE buffer – 1000ml: 10×TBE 50ml + H_2O to 1000ml.

10×Protein Gel Loading Buffer – 10ml final Con., stored at -20°C

1M Tris.Cl(pH7.5)	2.5ml	250mM
Bromophenol blue	20mg	0.2%
Glycerol	4ml	40%
H_2O	3.5ml	

Protein Lysis buffer - 10ml

1M Tris.Cl(pH8.0)	200ul
0.5M EDTA(pH8.0)	100ul
0.5% Triton X-100	50ul
Protease inhibitor	400ul (Stock: one tablet in 2ml PBS)
H_2O	9.25ml

4% Polyacrylamide Gel preparation—20ml

Distilled water	16.2ml
10×TBE buffer	1.0ml
37.5:1 acrylamide/bisacrylamide(40%)	1.25ml
40% acrylamide (w/v)	0.75ml
80% glycerol	625µl
TEMED	10µl
10%APS	150µl – added before use

80% glycerol – 10ml: Glycerol 8ml + H_2O 2ml

10% APS – 10ml, stored at -20°C: APS (ammonium persulfate) 1g + H_2O to 10ml.

40% acrylamide – 10ml, stored at 4°C: Acrylamide 4g + H_2O to 10ml.

Oligo diluting buffer (pH8.0) – 10ml final concentration

1M Tris (pH8.0)	100µl	10mM Tris
0.5M EDTA(pH8.0)	20µl	1mM EDTA
1M NaCl	500µl	50mM NaCl
H_2O	9380µl	

10% FCS-DMEM cell culture medium

DMEM

10% FCS (Fetal Calf Serum)

1% Penicillin/Streptomycin

10% FCS-RPMI 1640 cell culture medium

RPMI 1640

10% FCS (Fetal Calf Serum)

1% Penicillin/Streptomycin

12% Western Blot Resolving Gels-10ml

H_2O	3.3ml
30%Acrylamide mix	4ml
1.5M Tris(pH8.8)	2.5ml
10% SDS	100ul
TEMED	4ul
10% APS	100ul – added before use

4×Western Blot Electrophoresis buffer-1L

Tris 12g + Glycine 57.6g + SDS 4g or 10% SDS 200ml + H_2O to 1L.

1×Western Blot Electrophoresis buffer-1L: 4×Electrophoresis buffer 250ml + H_2O to 1L.

10×Western Blot Blotting buffer-1L

Tris 3g + Glycine (C2H5NO2) 14.4g + ethanol 100ml + H_2O to 1L.

1×Western Blot Blotting buffer-1L: 10×Blot buffer 100ml + methanol 100ml + H_2O to 1L.

5% Western Blot Stacking Gels-5ml

H2O	3.4ml
30%Acrylamide mix	0.83ml
1M Tris(pH6.8)	0.63ml
10% SDS	50ul
TEMED	5ul
10% APS	50ul – added before use

2×SDS Protein loading buffer-10ml

10%SDS	4ml
glycerol	2ml
0.5M Tris.Cl(pH6.8)	2ml
bromophenol blue	20mg or little
H_2O	2ml

2% milk buffer-500ml: 10g Protein88 + PBS 500ml, mix well and store at -4°C.

2.1.10 Primers

1740bp-KpnI Fw	5'- tag gta ccc ctg cca cag ggc aag aca g -3'
1740bp-XhoI Re	5'- tac tcg agg cag ccg gat gag gtg ctc g -3'
421bp-KpnI Fw	5'- tag gta ccc ttc tgc ctt ctc tgc cct ctc tt -3'
304bp-KpnI Fw	5'- tag gta ccc tca tat ata ctc aag gtt tcc -3'
250bp-KpnI Fw	5'- tag gta cca tgg gag tgg agg gcc acc aa -3'
216bp-KpnI Fw	5'- tag gta cca tgc gcg gaa gtc gaa ggc g -3'
176bp-KpnI Fw	5'- tag gta cct cca cgg ccg ccc cgt acc gtc c -3'
162bp-KpnI Fw	5'- tag gta ccg tac cgt ccg ggc agc gcg tt -3'
79bp-KpnI Fw	5'- tag gta ccG GCG AGC GGC GGG TGT CT -3'
Xho I Re	5'- tac tcg agg cag ccg gat gag gtg ctc gt -3'
5'-pho 40bp Fw	5'- atg cgc gga agt cga agg cgg cgt ctt agg gtg ttt tgg g -3'
5'-pho 40bp Re	5'- ccc aaa aca ccc taa gac gcc gcc ttc gac ttc cgc gca t -3'
75/63/54-KpnI Fw	5'- tag gta cca tgc gcg gaa gtc gaa ggc g -3' (216bp-KpnI Fw)
75-XhoI Re	5'- tac tcg aga acg cgc tgc ccg gac ggt ac -3'

63-XhoI Re	5’- tac tcg agg gac ggt acg ggg cgg ccg tgg a -3’
54-XhoI Re	5’- tac tcg agg ggg cgg ccg tgg acc caa aac a -3’
61bp-KpnI Fw	5’- tag gta ccg cgc gga agt cga agg cgg cg -3’
60bp-KpnI Fw	5’- tag gta ccc gcg gaa gtc gaa ggc ggc g -3’
50bp-KpnI Fw	5’- tag gta ccg aag gcg gcg tct tag ggt g -3’
61/60/50-XhoI Re	5’- tac tcg agg gac ggt acg ggg cgg ccg tgg a -3’ (63-XhoI Re)
Sp1-111Mu-KpnI Fw	5’- tag gta cca tgc gcg gaa gtc gaa Tgc gg -3’
Sp1-121Mu-KpnI Fw	5’- tag gta cca tgc gcg gaa gtc gaa ggc ggT -3’
Sp1-1Mu-XhoI Re	5’- tac tcg agg gac ggt acg ggg cgg ccg tgg a -3’ (63-XhoI Re)
RAP1 T-Mu-KpnI Fw	5’- tag gta cca tgc gcg gaa gtc gaa ggc ggc gtc tta ggT -3’
RAP1 T-Mu-XhoI Re	5’- tac tcg agg gac ggt acg ggg cgg ccg tgg a -3’ (63-XhoI Re)
Sp1-2Mu-KpnI Fw	5’- tag gta cca tgc gcg gaa gtc gaa ggc g -3’ (216bp-KpnI Fw)
Sp1-21Mu-XhoI Re	5’- tac tcg agg gac ggt acA ggg cgg ccg tgg a -3’
Oct-1-1Mu-KpnI Fw	5’- tag gta ccc tca tat ata Ttc aag gtt tcc -3’
Oct-1-2Mu-KpnI Fw	5’- tag gta ccc tca tat atG Ttc aag gtt tcc -3’
Oct-1-XhoI Re	5’- tac tcg agg cag ccg gat gag gtg ctc gt -3’ (Xho I Re)
Sp1-1 Bio Fw	5’- CGC GGA AGT CGA AGG CGG CGT CTT A -3’
Sp1-1 Bio Re	5’- TAA GAC GCC GCC TTC GAC TTC CGC G -3’
Sp1-1 n- Fw	5’- CGC GGA AGT CGA AGG CGG CGT CTT A -3’
Sp1-1 n- Re	5’- TAA GAC GCC GCC TTC GAC TTC CGC G -3’
RAP1 Bio Fw	5’- CGT CTT AGG GTG TTT TGG GTC CAC G -3’
RAP1 Bio Re	5’- CGT GGA CCC AAA ACA CCC TAA GAC G -3’
RAP1 n- Fw	5’- CGT CTT AGG GTG TTT TGG GTC CAC G -3’
RAP1 n- Re	5’- CGT GGA CCC AAA ACA CCC TAA GAC G -3’
Sp1-2 Bio Fw	5’- GGT CCA CGG CCG CCC CGT ACC GTC C -3’
Sp1-2 Bio Re	5’- GGA CGG TAC GGG GCG GCC GTG GAC C -3’
Sp1-2 n- Fw	5’- GGT CCA CGG CCG CCC CGT ACC GTC C -3’
Sp1-2 n- Re	5’- GGA CGG TAC GGG GCG GCC GTG GAC C -3’
Oct-1 Bio Fw	5’- AAC TCT CTC ATA TAT ACT CAA -3’
Oct-1 Bio Re	5’- TTG AGT ATA TAT GAG AGA GTT -3’
Oct-1 n-Fw	5’- AAC TCT CTC ATA TAT ACT CAA -3’
Oct-1 n-Re	5’- TTG AGT ATA TAT GAG AGA GTT -3’
SLA/LP 1512bp Fw	5’- TA GGA TCC ATG AAC CCG GAG AGC TTC GC-3’
SLA/LP 1512bp Re	5’- AA GAA TTC TAG AGC AGG GCC CTG GCC CA-3’

2.1.11 Laboratory instruments

BioPhotometer	Eppendorf AG, Hamburg, Germany
BioTrace PVDF membrane	Pall Corporation
CO2-AUTO-ZERO Incubator	Heraeus
Centrifuge 5417R	Eppendorf, Hamburg, Germany
DNA Engine Dyad Peltier Thermal	Cycler MJ Research, MC, USA
FastPrep Instrument	BIO 101 inc. CA, USA
Gel Doc 2000 System	Bio-Rad Laboratories GmbH, München, Germany
Hybond-N+ membrane	Amersham, #RPN203B
HL-2000 HybriLinker	UVP, Inc., USA
inoLab pH Level 1	WTW, Weilheim, Germany
Kodak BioMax MR-1 films	Integra Biosciences GmbH, Fernwald, Germany
Lumat LB 9507 Tube Luminometer	Berthold Technologies GmbH, Bad Wildbad, Germany
Lysing-Matrix-D tubes	BIO 101 inc., CA, USA
Magnetic mixer, MR3001	Heidolph Instruments GmbH. Nürnberg, Germany
Mini-Protein 3 Electrophoresis Cell	Bio-Rab Laboratories GmbH, München, Germany
Mini Trans-Blot Transfer Cell	Bio-Rad Laboratories GmbH, München, Germany
Rocking, Duomax 1030	Heidolph Instruments GmbH. Nürnberg, Germany
Roller Mixer, SRT1	Barloworld Scientific, UK
Storm 860 scanner	GE Healthcare, Freiburg, Germany
Thermomixer Comfort	Eppendorf, Hamburg, Germany
TMS-F Microscope	Nikon, Japan
-85℃ Ultralow Freezer	NUAIR, USA

2.1.12 Prediction Programs

A promoter scan program, WWW Promoter Scan, was used to predict SLA/LP gene promoter. The web site is http://bimas.dcrt.nih.gov/molbio/proscan/.

A web-based transcription factor binding site identification program, AliBaba 2.1 (Grabe 2002), was use to predict SLA/LP gene transcription factor binding sites. The web site is http://www.gene-regulation.com/pub/programs.html.

A CpG island promoter detection algorithm, CpGProD (Ponger et al. 2001), http://pbil.univ-lyon1.fr/software/cpgprod_query.php), was used to predict the possibility of a CpG island promoter.

2.2 Methods

2.2.1 Preparation of plasmid DNA

All the plasmid DNAs were prepared with QIAprep Spin Miniprep Kit or HiSpeed Plasmid Maxi Kit (Qiagen) according to the manuals. DNA was dissolved in H_2O when used immediately; otherwise, the DNA was dissolved in low TE (pH8.0) buffer or 10mM Tris.HCl (pH8.5).

2.2.2 Polymerase chain reaction (PCR)

* PCR with plasmid DNA – REDtaq ReadyMix PCR Kit (Sigma)

In general, the final concentration of primers was 1µM and 0.5µl of plasmid DNA in a 25µl volume reaction. The standard protocol was as following:

- Initial denaturation at 95°C for 5 min, then 28 – 35 cycles of:
  - Denaturation 45s at 95°C
  - Annealing 45s at 55-68°C [depending on the melting temperature (Tm) of the primers]
  - Elongation 45s-4min at 72°C (depending on fragment length: 45s-0.75kb; 1min-1.5kb; 2min-3kb; 4min-6kb; 8min-10kb; 68°C if the fragment > 3kb)
- Final elongation 5min at 72°C (68°C if the fragment larger than 3kb)

* PCR with bacteria solution - REDtaq ReadyMix PCR Kit (Sigma)

Some bacteria were picked from a clone and suspended in 10µl H_2O; 2-5µl of this solution was used in 15µl of PCR.

The protocol was the same as that for PCR with plasmid DNA, but the initial denaturation time was changed to 10min.

* High Fidelity PCR – Expand High Fidelity PCR System (Roche)

The purpose of this PCR was to get high fidelity and high specificity PCR products because the Tgo DNA polymerase is a thermostable DNA polymerase with proofreading activity. The reaction was performed according to the protocol of the kit.

2.2.3 DNA electrophoresis

Agarose gels of 0.75% to 1.0% were used for a wide range of separations (0.5 to 15 kb). 2-4% agarose gels were usually selected for separating small PCR fragments or separating fragments that only had several bps differences. See Table 1.

Table 1 Relationship between Agarose Gel and DNA Size

% agarose	DNA (bp)
0.75	10 000 – 15 000
1.0	500 – 10 000
1.25	300 – 5 000
1.5	200 – 4 000
2	100 – 2 500
2.5	50 – 1 000

1×TAE buffer provided optimal resolution of fragments > 4 kb in length, while for 0.1 to 3 kb fragments 0.5×TBE buffer was selected. TBE had both a higher buffering capacity and a lower conductivity than TAE and therefore was used for high voltage electrophoresis.

2.2.4 DNA extraction and purification

* **DNA extraction from agarose gel** -- QIAquick Gel Extraction Kit (Qiagen)

300mg of DNA agarose gel – Add 900µl of Buffer QG – Incubate 10min at 50°C to melt the gel – Add 300µl of Isopropanol, mix – Load the solution to a Spin column (750µl/time) – 13000rpm for 30s – Wash the column with 750µl of Buffer PE -- 13000rpm for 30s – Discard supernatant and 13000rpm for 1min – Add 30µl H_2O to the center of the column – Incubate 3min at RT – 13000rpm for 1min, collect the passthrough – Test the DNA concentration.

* **DNA extraction and purification from enzyme digestion** – Phenol-Chloroform method

- 0.5ml sample + 0.5ml Phenol/Chloroform/Isoamyl alcohol (25:24:1) (Roth, #A156.1), mix, 13000 rpm for 2min.
- Transfer the aqueous (upper) phase to a fresh tube; add 0.5ml Phenol/Chloroform/Isoamyl alcohol, mix, 13000rpm, 2min.
- Transfer the upper phase to a fresh tube; add 0.5ml Chloroform/Isoamyl alcohol (24:1), mix, 13000 rpm for 2min.
- Transfer the upper phase to a fresh tube, add 1/10V 8M LiCl, mix; add 0.8V Isopropanol, mix, -70˚C, 1hour.
- 4˚C, 13000rpm for 20min. Discard the supernatant.
- Wash the pellet with 70% EtOH, 4˚C, 13000rpm, 5min × 2

- Dry the pellet in air and redissolve the pellet with H_2O.

2.2.5 Northern Blot

2.2.5.1 RNA isolation

Liver sample (100mg) from an 8 weeks old C57BL/6 mouse was homogenized in a Lysing-Matrix-D tube using 1ml TriReagent according to the manual instructions. RNA was dissolved in RNase free water, measured with GeneQuantpro and frozen at -80°C.

2.2.5.2 Reverse Transcription (RT) and SLA/LP mRNA-specific PCR (RT-PCR)

Reverse transcription of 0.5µg of RNA was performed with the ProSTAR First-Strand RT-PCR Kit according to the manufacturer. The cDNA was used as template to get a 1512bp SLA/LP fragment with the primers, SLA/LP 1512bp Fw and SLA/LP 1512bp Re. The PCR was conducted with annealing temperature at 58°C, elongation time for 2min and 30 cycles. The PCR product was gel purified and digested with BamH I/EcoR I, then subcloned into BamH I/EcoR I site of pET-30a (+) vector. This clone was used for preparation of Northern Blot probe.

2.2.5.3 Preparation of Northern Blot probe

The 1512bp/pET-30a (+) clone was digested with Pst I. The 448bp fragment was extracted from the agarose gel and subcloned into pBlueScript SK (+) vector digested with Pst I. The clone was sequenced and used to cut out the 448bp probe for hybridization.

2.2.5.4 Labeling the Northern Blot probe with isotope

The 448bp probe (30ng) was α-^{32}P-dCTP labeled with Prime-It II Random Primer Labelling Kit according to the manual and purified with ProbeQuant G-50 micro column.

2.2.5.5 Northern Blot of Mouse Embryo and mouse tissues

The membranes loaded with mouse embryonic mRNA or mRNA from various tissues of adult mice were purchased from Clontech or OriGene. The Northern Blot with the 488bp α-^{32}P-dCTP labeled SLA/LP probe was performed according to the manufacturer. The membrane was exposed to phosphoimager film overnight and scanned with Storm 860 scanner (GE Healthcare, Freiburg).

2.2.6 Western Blot

2.2.6.1 Isolation of protein from different tissues

A little liquid nitrogen was added to 100mg of tissue, which was crushed with a mortar, and transferred to a 1.5ml tube. 300ul of Protein Lysis Buffer was added and the mixture was incubated for 30min at 4°C. Then, the mixture was centrifuged for 2min at 6000rpm and 4°C, and the supernatant (protein mixture) was transferred to a fresh tube. The concentration of protein was measured with Bio-Rad Protein Assay according to the manual.

2.2.6.2 Western Blot

Proteins were separated by electrophoresis on a 12% SDS-Polyacrylamide gel. Briefly, the same amount of proteins from different tissues was loaded with 2×SDS Protein Loading Buffer onto a 5% stacking gel, and separated through a 12% SDS-resolving gel by vertical electrophoresis for 75min at 150 Volts.

Proteins were blotted to Polyvinylidene fluoride (PVDF) membrane by electrotransfer. Briefly, PVDF membrane was equilibrated first with 100% methanol for 2min, and then with 1×Blotting Buffer for 15min. The gel was washed in 1×Blotting Buffer for 10min. Proteins were then transferred to membrane for 1h at 500V/150mA.

The membrane was blocked in 2% milk buffer for 1 hour at room temperature, incubated overnight at 4°C with Anti-SLA/LP positive human serum, which did not contain reactions to other known autoantibodies (1:200 in milk buffer). After washing with PBS-0.1%Tween 20, bound SLA/LP antibody, was detected by incubating the membrane with anti-human HRP-IgA/G/M (1:1000) (DAKO) for 1h at RT, followed by DAB staining.

2.2.7 Cloning and sequencing of mouse SLA/LP promoter fragment

2.2.7.1 Sticky-end cloning into plasmid vector

The general process was as following:

Digestion – Electrophoresis – Extraction – Ligation – Transformation – Selection - Sequencing.

- Digestion: 1unit of enzyme digests 1μg of DNA in 1 hour at 37°C. Some enzymes have star activity, that restriction endonucleases are capable of cleaving sequences which are similar but not identical to their defined recognition sequence under extreme unstandard condition, such as BamH I, EcoR I and Sal I. Star activity can be

reduced by limiting amount of enzyme and digestion time.

- Electrophoresis: Digestion products were separated by agarose gel electrophoresis.
- Extraction: DNA was recovered from the gel using QIAquick Gel Extraction Kit (Qiagen).
- Ligation: The ratio of insert to vector was between 3~5 to 1. The amount of insert was calculated according to the formula, ng of insert = (100ng of vector × kb of insert / kb of vector) × 3 ~5/1 (Promega manual, #TM042). Ligation was performed with the Quick Ligation Kit (New England BioLabs).
- Transformation: DNA was less than 50ng for 100ul of competent cells. The pGL3 Basic vector contains ampicillin resistance. The XL2-Blue or XL10-Gold Ultracompetent cells (Stratagene) were used for transformation according to the manual.
- Selection: Transformation was controlled by PCR or by restriction productions. PCR selection was easier and faster than digestion when there were many clones to be selected.

2.2.7.2 Blunt-end cloning into plasmid vector

The general process was as following:

Digestion – Blunt – Dephosphorylating vector – Electrophoresis – Extraction – Ligation – Transformation – Selection.

- Blunt: 1unit of Klenow enzyme was used for blunting 1μg of DNA; dNTP concentration was 33μM in final. A 5'- end overhang DNA was easier to be blunted (fill-in) than a 3'-end overhang DNA (chew-back).
- Dephosphorylation of vector: Different alkaline phosphatase, calf intestinal alkaline phosphatase (CIP), shrimp alkaline phosphatase (SAP) and antarctic phosphatase, was used to dephosphorylate vector. Antarctic phosphatase was chosen to dephosphorylate vector because it was more efficient. When the vector DNA was blunt end or 3' end extension, dephosphorylation was performed at 50°C and with more enzymes.
- Ligation: The ratio of insert to vector was ranging from 3:1 to 10:1 according to their size.

2.2.8 The generation of 1740bp fragment clone

A 1740bp DNA fragment (-1623 to +117bp) (The accession number is NM-172490) upstream of the translation start site of the SLA/LP was PCR amplified with High Fidelity PCR Kit (Roche) which had 3' to 5' proofreading activity. The primers were 1758bp-KpnI

Fw and 1758bp-XhoI Re, which were designed based on the information from the mouse BAC library database describing the sequence of mouse chromosome 5. The SLA/LP BAC DNA (Clontech) was used as the template. The PCR product was gel purified and double digested with Kpn I /Xho I. The luciferase gene reporter vector, pGL3 Basic vector, was also digested with Kpn I/Xho I, and then ligated with the digested PCR fragment and transformed into XL-2 Blue competent cells. The positive clone was selected by PCR and sequenced.

The 1740bp fragment was subcloned into pGL3 Basic vector with two directions, 5' to 3' and 3' to 5' direction. The purpose was to check whether the initiating ability was orientation dependent.

The cloning procedure is shown in Fig 6.

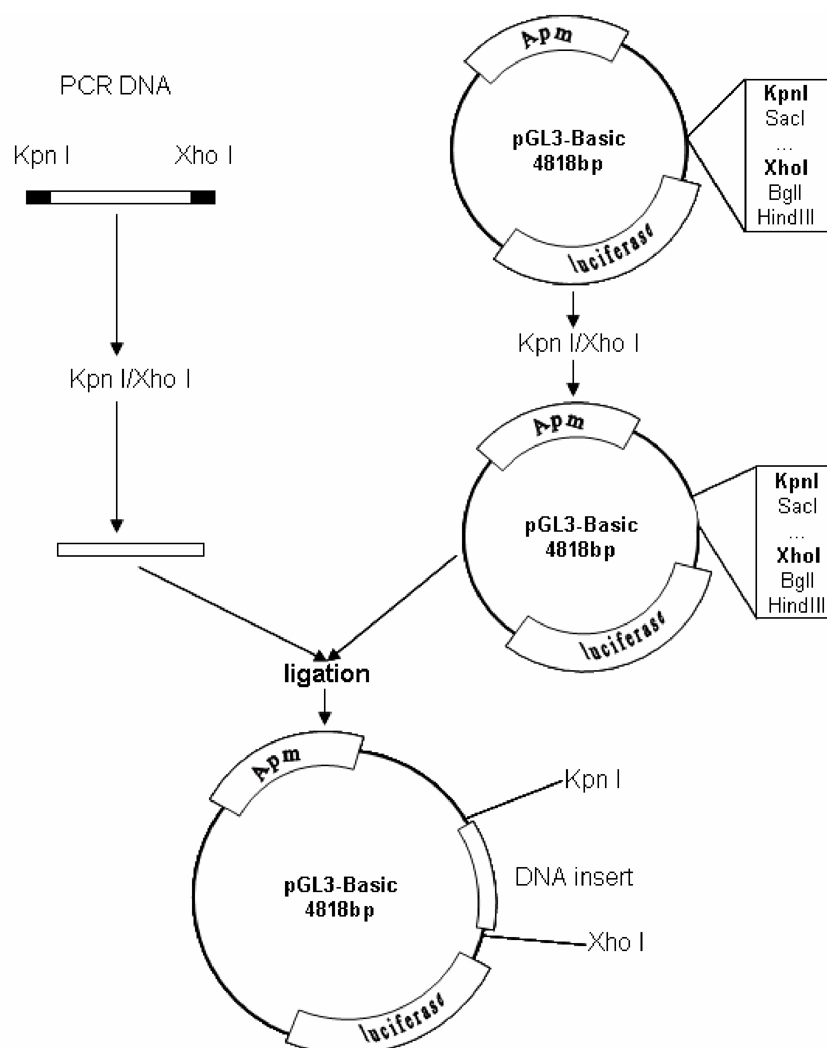

Fig 6. The construction of different SLA/LP gene clones. The PCR primers contained KpnI site in forward prime and XhoI site in reverse primer. After digested with KpnI/XhoI, the PCR fragment was cloned into pGL3 Basic vector digested with the same enzymes.

2.2.9 5'-end deletion and 3'-end deletion

After confirming the initiating transcription ability of the 1740bp fragment by Luciferase Assay, this clone DNA was digested with different enzymes (XmaI, EcoRI, SauI and SpeI), then self-ligated to synthesize serial 5'-end deletion mutations (1251bp, 749bp, 570bp and 391bp) and tested for luciferase activity. According to the results, further 5'-end deletions were conducted. The mutants were in sizes of 421bp, 391bp, 304bp, 250bp, 216bp, 176bp, 162bp and 79bp, which were obtained by PCR with primers containing Kpn I and Xho I restriction sites.

The 3'-end deletions were performed based on the results of 5'-deletions. The 216bp mutant, which might contain the SLA/LP promoter and possible transcription factor binding sites, was used for 3'-end deletion. The 3'-end deletion fragments, 75bp, 63bp and 54bp, were obtained by serial PCRs with the same primers as above, and subcloned into pGL3 Basic vectors separately. They were used for further luciferase assays after sequencing.

2.2.10 Point mutation in the sequence of mouse SLA/LP promoter

The minimal stretch of SLA/LP DNA fragment, which could initiate transcription, was identified after serial deletions. Point mutations then were introduced into the possible transcription factor binding sites by PCR. The primers were Sp1-111Mu-KpnI Fw, Sp1-121Mu-KpnI Fw and Sp1-1Mu-XhoI Re for the mutations of the Sp1 site at -85 to -76. RAP1 T-Mu-KpnI Fw and RAP1 T-Mu-XhoI Re were used for mutating the putative RAP1 site at -71 to -62. Sp1-2Mu-KpnI Fw and Sp1-21Mu-XhoI Re were used for mutating the putative second Sp1 site at -55 to -41. Finally, Oct-1-1Mu-KpnI Fw, Oct-1-2Mu-KpnI Fw and Oct-1-XhoI Re were used for the point mutations of presumed Oct-1 site at -184 to -175. The primers are shown in the list of primers in the materials section. The PCR protocols were adapted according to the Tm of each primer pair. The PCR products were KpnI/XhoI digested, gel purified and subcloned into KpnI/XhoI site of pGL3 Basic vector, sequenced and used for luciferase assays to show the changes after mutations.

2.2.11 Transient transfection

Transient transfection of HEK293, RAW264.7 and Hepa1-6 cell lines with plasmids were performed by FuGENE6 transfection reagent (Roche). All transfections were performed with internal control by co-transfection with β-glactosidase plasmid. Luciferase activity was expressed as relative activity, calculated by dividing specific luciferase units by β-glactosidase units. The purpose of using three cell lines was to identify cell-specific patterns.

Briefly, the transient transfection was performed by mixing 3µl of FuGENE6 transfection reagent with 97µl of RPMI1640 or DMEM without serum and antibiotics, followed by 5 minutes incubation at room temperature, and addition of 1µg of SLA/LP fragment/pGL3 Basic vector constructs and 0.7µg of β-glactosidase plasmid, followed by another 20 minutes of incubation at room temperature. Then, the transfection mixture was added drop-wise to 24 hours cultures of 2.2×10^5 Hepa1-6 cells or 1.8×10^5 HEK293 cells in 2ml of 10%FCS-DMEM, or to 1.8×10^5 RAW264.7 cells in 2ml of 10%FCS-RPMI1640, each plated in 6 well-plates, followed by 48 hours of incubation in a humidified incubator equilibrated with 5% CO_2 at 37°C.

2.2.12 Luciferase assay

*** Preparation of cell lysates**

After incubation for 48 hours post-transfection, the cells were washed twice with PBS and lysed with 230µl of Reporter Lysis Buffer (RLB, Promega) for 15 min at room temperature. All cells were scraped from the wells and transferred to 1.5ml tubes. The cell lysates were vortexed for 15 seconds, centrifuged for 2min at 13000rpm and 4°C. The supernatants were collected and stored at -70°C.

*** Luciferase Assay**

Twenty microliters of cell lysate was transferred to a tube and firefly luciferase activity was measured with 100µl of luciferase substrate (Luciferase Assay System, Promega) and a Lumat LB 9507 Tube Luminometer (Berthold, Germany). All samples were tested twice. The luminometer was programmed to perform a 2-sec pre-measurement delay, followed by a 10-sec measurement period for each reporter assay. One hundred microliters of luciferase substrate were automatically injected and firefly luciferase activity was measured.

*** β-Galactosidase Assay**

The cell lysate was diluted 2:1 with RLB buffer (100µl of lysate + 50µl of RLB buffer), 150 µl of 2×Assay Buffer was added, mixed and incubated at 37°C for 30min or until a faint yellow colour had developed (30min for HEK293 cell line and 4h for RAW264.7 and Hepa1-6 cell line.). The reaction was stopped by adding 500 µl of 1M NaOH and the absorbance was read at 420nm. The β-Galactosidase units were obtained by comparing with the standard curve.

*** The relative luciferase activities units (RLU)**

The firefly luciferase activity of each sample was normalized to β-glactosidase activity and each sample was then compared to the maximal activity.

Each assay was repeated at least three times.

2.2.13 Gel Shift Assay and Super Gel Shift Assay

Gel shift assay, or electrophoretic mobility shift assay (EMSA), has been used widely in the study of sequence-specific DNA-binding proteins such as transcription factors. The assay is based on the observation that complexes of protein and DNA migrate through a nondenaturing polyacrylamide gel more slowly than free DNA fragments or double-stranded oligonucleotides. Thus, if the labelled putative transcription factor binding oligonucleotides can form a complex with the factor protein, there will be a shifted band appearing whose position is higher than the free probes. The specificity of the DNA-binding protein for the putative binding site is established by competition experiments using DNA fragments or oligonucleotides containing a binding site for the protein of interest. The specificity also can be further confirmed by super gel shift. Super gel shift need incubate the gel shift reaction system with antibody and will appear a super shifted band which position is higher than the shifted band if the labelled Oligonucleotides-factor protein-antibody complex forms.

The gel shift assays and super gel shift assays were performed with Biotin-labelled probes. The DNA binding reactions were conducted with the Gel Shift Assay System from Promega, while the detections were performed with LightShift Chemiluminescent EMSA Kit from PIERCE.

2.2.13.1 Preparation of double-stranded DNA probes

The Biotin-labeled probes and unlabeled probes were diluted to 20fmol/µl or 4pmol/µl with the Oligo diluting buffer (pH8.0). The dilution was prepared freshly because oligonucleotides were not stable at low concentration.

The preparation of double stranded DNA was performed by mixing the same amount of Fw and Re oligos, incubation at 95°C for 5min and allowing to cool down slowly to room temperature. The double-stranded DNA probes were stored at -20°C. The probes were not refrozen and rethawed for more than two cycles.

2.2.13.2 DNA Binding Reactions

The conditions of DNA binding reactions were adjusted according to the studied DNA elements:

- For testing the possible Sp1 site at -85 to -76, the recombinant human Sp1 protein (rhSp1, Upstate) was used for binding Sp1 consensus sequence. The reaction is shown in Table 2.

10×gel loading buffer was added only to the negative control sample because Sp1 transcription factor is sensitive to dyes (Promega manual, TB110).

Table 2 Gel Shift Reaction System of the Putative Sp1 Site at -85 to -76

	Neg.	Sample	Competitor
Nuclease-Free H_2O	7.3µl	4.8µl	2.8µl
5×Gel Shift Binding Buffer	2.0µl	2.0µl	2.0µl
171ng/µl rhSp1	---	2.5µl	2.5µl
4pmol/µl Unlabeled competitor oligos	---	---	2.0µl
Incubate 10min at RT(23°C) to overcome strong nonspecific interactions			
20fmol/µl Labeled consensus oligos	0.7µl	0.7µl	0.7µl
Incubate 20min at RT (23°C)			
10× protein gel loading buffer	1.0µl	---	---

- For testing the second possible Sp1 site at -55 to -41, the recombinant human Sp1 protein (rhSp1, Upstate) was used. The reaction is shown in Table 3.

Table 3 Gel Shift Reaction System of the Putative Sp1 Site at -55 to -41

	Neg.	Sample	Competitor
Nuclease-Free H_2O	7.5µl	5.0µl	2.0µl
5×Gel Shift Binding Buffer	2.0µl	2.0µl	2.0µl
171ng/µl rhSp1	---	2.5µl	2.5µl
4pmol/µl Unlabeled competitor oligos	---	---	3.0µl
Incubate 10min at RT(23°C) to overcome strong nonspecific interactions			
20fmol/µl Labeled consensus oligos	0.5µl	0.5µl	0.5µl
Incubate 20min at RT (23°C)			
10× protein gel loading buffer	1.0µl	---	---

- For testing the presumed RAP1 (-71 to -62) and Oct-1 (-184 to -175) binding site, the HeLa nuclear extract (Promega) was used. The DNA binding reaction condition of RAP1 or Oct-1 is shown in Table 4.

Table 4 Gel Shift Reaction System of the Putative RAP1 or Oct-1 Site

	Neg.	Sample	Competitor
Nuclease-Free H_2O	7.5µl	6.5µl	3.5µl
5×Gel Shift Binding Buffer	2.0µl	2.0µl	2.0µl
HeLa nuclear extract	---	1.0µl	1.0µl
4pmol/µl Unlabeled competitor oligos	---	---	3.0µl
Incubate 10min at RT(23°C) to overcome strong nonspecific interactions			
20fmol/µl Labeled consensus oligos	0.5µl	0.5µl	0.5µl
Incubate 20min at RT (23°C)			
10× protein gel loading buffer	1.0µl	---	---

2.2.13.3 Super Gel Shift Assay

For Super Gel Shift Assay of Sp1 site, 1.5µl of 1µg/µl Anti-Sp1 antibodies (Upstate, EMSA tested) were added to the reaction system after the 20min incubation and incubated for further 30min at room temperature.

For Super Gel Shift Assay of RAP1 site, 6µl of 0.7634µg/µl Anti-RAP1 antibodies (Upstate, not EMSA tested) were added to the reaction system after the 20min incubation and incubated for further 30min at room temperature.

For Super Gel Shift Assay of Oct-1 site, 1.0µl of mouse ascites containing Anti-Oct-1 antibodies (Upstate, EMSA tested) were added to the reaction system after the 20min incubation and incubated for further 30min at room temperature.

2.2.13.4 Electrophoresis

The gel shift and super gel shift reactions were separated on 4% Polyacrylamide gel by electrophoresis. The preparation of the gel is shown in the materials section. The gel did not contain SDS, which was used in normal SDS-PAGE, because some transcription factors were very sensitive to detergents.

The 4% Polyacrylamide Gel was pre-run at 300V for 15min in 0.5×TBE buffer.

Samples were loaded onto the gel carefully. It was important that the 10× protein gel loading buffer was only added to the negative control when the tested factor was not mentioned to be sensitive to dye or not. Electrophoresis was performed at 100V/200mA until the bromophenol blue dye had migrated approximately 3/4 down the length of the gel in 0.5×TBE buffer. The gel temperature was maintained below 30°C.

2.2.13.5 Blotting

Hybond-N+ membrane (Amersham) was wetted with H_2O and incubated in 0.5×TBE buffer for 15min at room temperature with shaking.

The DNA/protein complexes were transferred to the membrane in a Mini Trans-Blot Chamber (Bio-Rad, München) by applying 100V/380mA in 0.5×TBE for 1 hour.

The membrane was placed with bromophenol blue side up on a paper towel and allowed to dry.

2.2.13.6 Cross-link Transferred DNA to membrane

The transferred DNA and DNA-protein were fixed to the membrane by UV light with UV-light cross-linker (HL-2000 HybriLinker, UVP): energy-120mJ/cm^2, time-60s. The membrane could be stored dry at room temperature for several days.

2.2.13.7 Detect Biotin-labeled DNA - Chemiluminescent Nucleic Acid Detection Module (Pierce)

- Slowly warm the Blocking Buffer and the 4×Wash Buffer to 37°C in a water bath until all particulates are dissolved.
- To block membrane, add 20 ml Blocking Buffer and incubate for 15 minutes with gentle shaking at room temperature.
- Prepare conjugate/blocking buffer solution by adding 66.7 µl of the Stabilized Streptavidin-Horseradish Peroxidase Conjugate to 20 ml Blocking Buffer (1:300 dilutions).
- Decant blocking buffer from the membrane and add 20 ml of the conjugate/blocking solution and incubate for 15minutes with gentle shaking.
- Prepare 1× wash solution by adding 30 ml of 4×Wash Buffer to 90 ml ultrapure water.
- Transfer membrane to a new container and rinse briefly with 20 ml of 1× wash solution.
- Wash membrane four times for 5min each in 20 ml of 1×wash solution with gentle shaking.
- Transfer membrane to a new container and add 30 ml of Substrate Equilibration Buffer. Incubate membrane for 5 minutes with gentle shaking.

- Prepare Chemiluminescent Substrate Working Solution by adding 6 ml Luminol/Enhancer Solution to 6 ml Stable Peroxide Solution. Avoid light.
- Remove membrane from the Substrate Equilibration Buffer and carefully blot an edge of the membrane on a paper towel to remove excess buffer. Place membrane in a clean container or onto a clean sheet of plastic wrap placed on a flat surface.
- Pour the Substrate Working Solution onto the membrane so that it completely covers the surface. Alternatively, the membrane may be placed nucleic acid side down onto a puddle of the Working Solution. Incubate membrane in the substrate solution for 5 minutes without shaking.
- Remove membrane from the Working Solution and blot an edge of the membrane on a paper towel for 2-5 seconds to remove excess buffer. Do not allow the membrane to become dry.
- Wrap the moist membrane in plastic wrap, avoiding bubbles and wrinkles.
- Place membrane in a film cassette and expose to X-ray film for 10 seconds to 5 minutes depending on the signals.

3. Results

3.1 Western Blot of different mouse tissues

To study the expression of SLA/LP protein on different tissues, Western Blot was conducted. Proteins extracted from thirteen different tissues of a 2 weeks old male B10.PL mouse were used to test the expression of SLA/LP protein. A patient serum solely positive for anti-SLA/LP antibodies was used for the detection. All tissues expressed SLA/LP protein to various extents. Here, pancreas showed the highest expression levels, followed by liver, testis, kidney and thymus; expression in muscle was the lowest one (Fig 7).

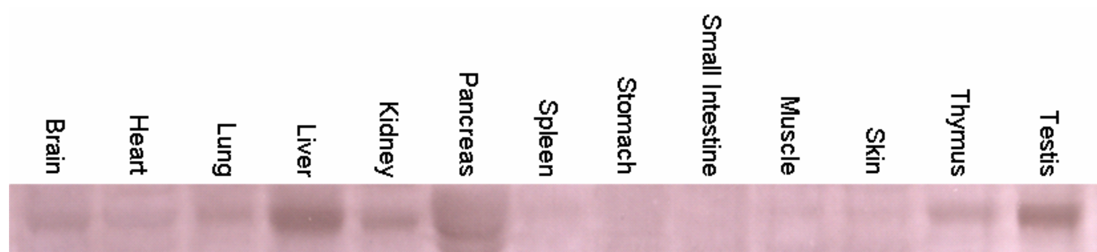

Fig 7. Western Blot of different mouse tissue proteins with Anti-SLA/LP positive human serum. SLA/LP protein was expressed widely in normal tissues.

3.2 Northern Blot on mouse embryos and different mouse tissues

SLA/LP protein was expressed widely in different tissues. To investigate when SLA/LP mRNA started to be transcripted and to confirm the extent of expression in different tissues, Northern Blots were performed with mouse embryonic RNA and RNA from adult tissues.

The mouse SLA/LP mRNA was faintly detectable from day 7 of embryonic development onward. Thus, SLA/LP protein might play a role during embryonic development (Fig 8).

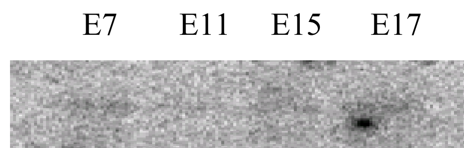

Fig 8. Northern Blot of mouse embryos with mouse SLA/LP-probe. As indicated, the mRNA, was derived from 7, 11, 15 or 17 days old mouse embryo. SLA/LP mRNA was faintly detectable from day 7 onward.

Northern Blot with RNA from different adult mouse tissues showed that SLA/LP mRNA was detectable in all tissues to various extents, the liver exhibiting highest and muscle exhibiting lowest expression. The ubiquitous expression of SLA/LP mRNA indicated that

SLA/LP gene might serve basic cellular functions. However, the high expression in liver suggests a more important role in liver function.

This membrane lacked pancreatic mRNA; we thus could not compare mRNA from liver and pancreas. See Fig 9.

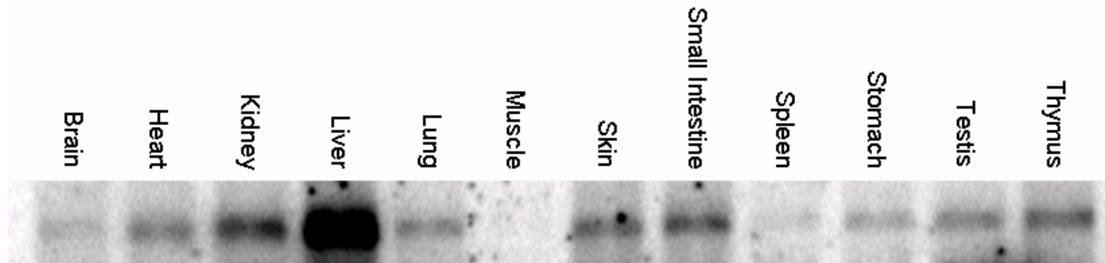

Fig 9. Northern Blot of different mouse tissues with mouse SLA/LP-probe. The SLA/LP mRNA was detected ubiquitously in normal tissues.

3.3 1740bp SLA/LP fragment and the ability to initiate protein expression

To locate the promoter region of SLA/LP protein, a 1740bp fragment (-1623 to +117) covering the translation initiation site of SLA/LP was cloned into pGL3 Basic vector in two directions and tested by Luciferase Assay. The 1740bp/pGL3 plasmids were transiently transfected into HEK293, RAW264.7 or Hepa1-6 cells. Compared to the vector control, relative luciferase activity of 1740bp/pGL3 Basic clone in 5' to 3' direction was very high, 117.9-fold higher in HEK293 cells, 166.2-fold higher in RAW264.7 cells and 21.7-fold higher in Hepa1-6 cells. Luciferase activity of the 1740bp/pGL3 Basic clone in 3' to 5' direction was negligible, ranging from 0.7 to 1.2-fold activity of vector control (Fig 10). These results indicate that the 1740bp fragment bears promoter function, which is orientation dependent.

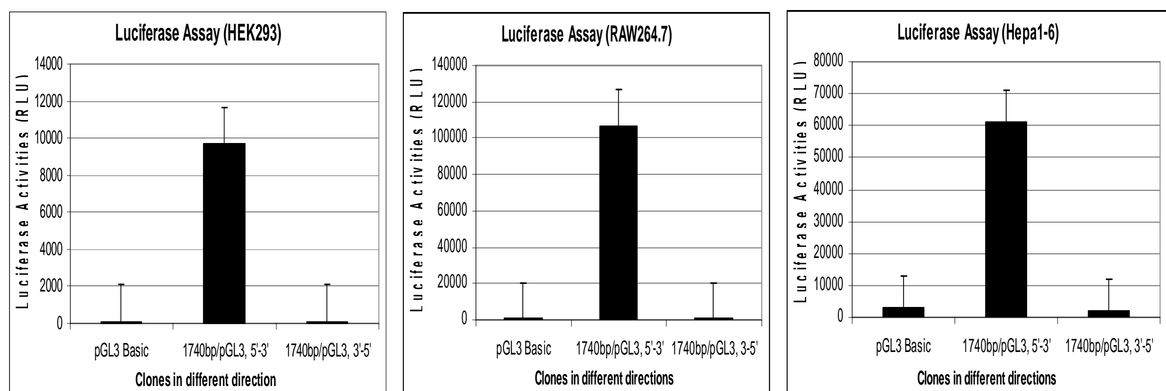

Fig 10. Luciferase assay of a 1740bp SLA/LP fragment in different orientations. Luciferase activity of 1740bp fragment in 5' to 3' direction was very high, while similar to empty vector control in 3' to 5' direction.

3.4 The typical core promoter elements in the 1740bp SLA/LP fragment

There are no typical core promoter elements in the 1740bp SLA/LP fragment, such as TATA box, DPE, BRE or Inr at the appropriate locations upstream of SLA/LP transcription start site. There is a TCAGTGG sequence motif, starting from -1 of SLA/LP transcription start site, reminiscent of the Inr consensus sequence (PyPyANTPyPy); however, the last two important pyrimidines (Javahery et al. 1994; Lo et al. 1996) are changed to purines.

3.5 5'- or 3'-deletion mutations of the 1740bp/pGL3 Basic clone and Luciferase Assay

To identify the core promoter sequence of the SLA/LP gene and possible transcription factor elements, 5'-end or 3'-end deletion mutations of the 1740bp fragment were performed. The DNA of the 1740bp/pGL3 Basic clone was digested by Xma I, EcoR I, Sau I or Spe I, self-religated and transformed into competent XL-2 Blue cells. This procedure yielded 5'-end deletion mutants of 1249bp, 749bp, 570bp or 391bp lengths. Moreover, 5'-end deletion mutants of 421bp, 304bp, 250bp, 216bp, 176bp, 162bp or 79bp lengths, and 3'-end deletion mutants of 75bp, 63bp or 54bp lengths were generated by PCR with primers containing Kpn I or Xho I site. These PCR fragments were digested with Kpn I and Xho I, subcloned into Kpn I/Xho I digested pGL3 Basic vector and used for transient transfection as described above.

The relative positions and lengths of these fragments are shown in Fig 11 and Fig 12.

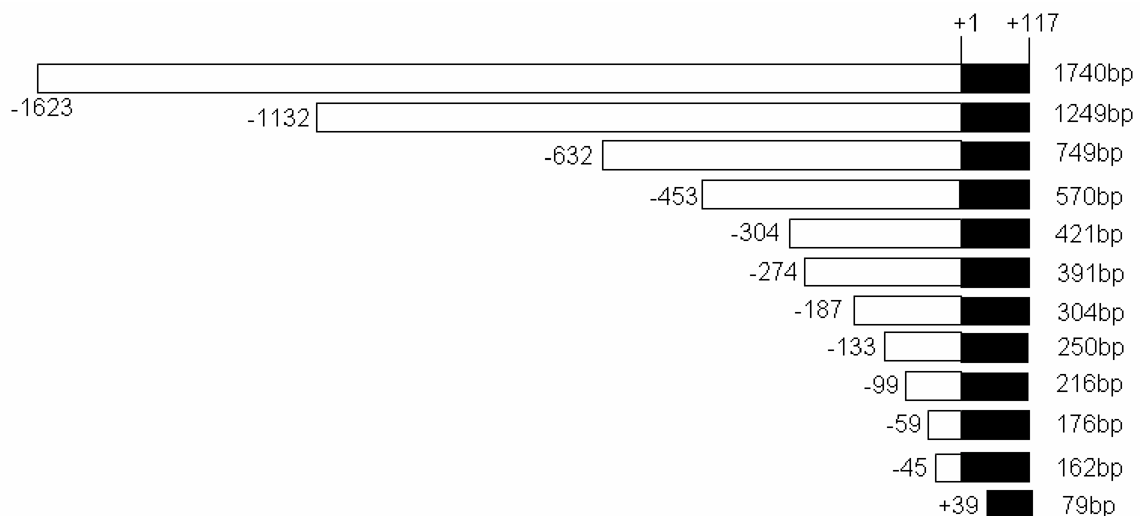

Fig 11. SLA/LP promoter variants generated by 5'-end deletion of the 1740bp SLA/LP promoter fragment. The length of the variant promoters was 1740bp, 1249bp, 749bp, 570bp, 421bp, 391bp, 304bp, 250bp, 216bp, 176bp, 162bp or 79bp.

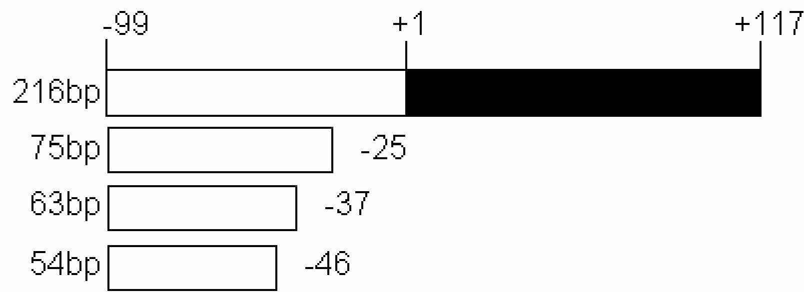

Fig 12. SLA/LP promoter variants generated by 3'-end deletion of the 216bp SLA/LP promoter fragment. The length of the variant promoters was 216bp, 75bp, 63bp or 54bp.

The luciferase assay results of these 5'-end deletion promoter fragments showed a similar pattern in the three studied cell lines (Fig 13A-C). There was a drop of luciferase activity in fragments shorter than 570bp, indicating that an activator sequence may be located between -453 and -274 upstream of the transcription start site. However, luciferase activity was greatly increased in the 250bp mutant, suggesting the existence of a repressor within the 54bp (-187 to -134) between fragment 304bp and 250bp. Further deletion resulting in a 176bp mutant was associated with a loss of luciferase activity, suggesting the presence of the SLA/LP core promoter within the 216bp mutant.

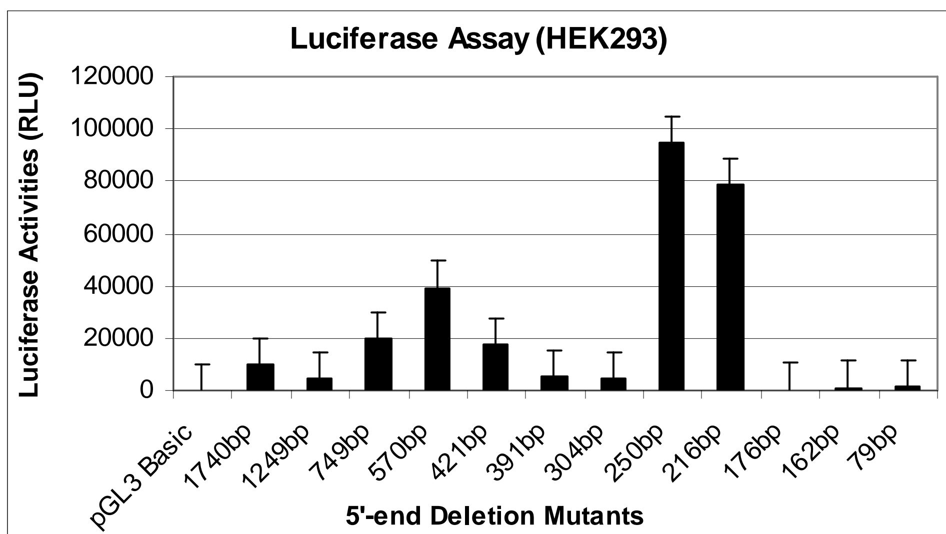

Fig 13A. Luciferase assay result of 5'-end deletion mutants in HEK293 cells.

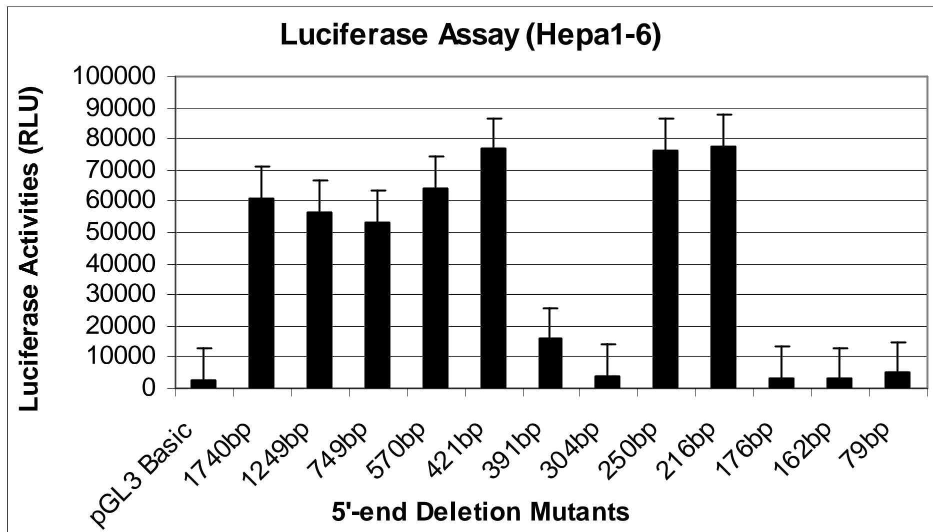

Fig 13B. Luciferase assay result of 5'-end deletion mutants in Hepa1-6 cells.

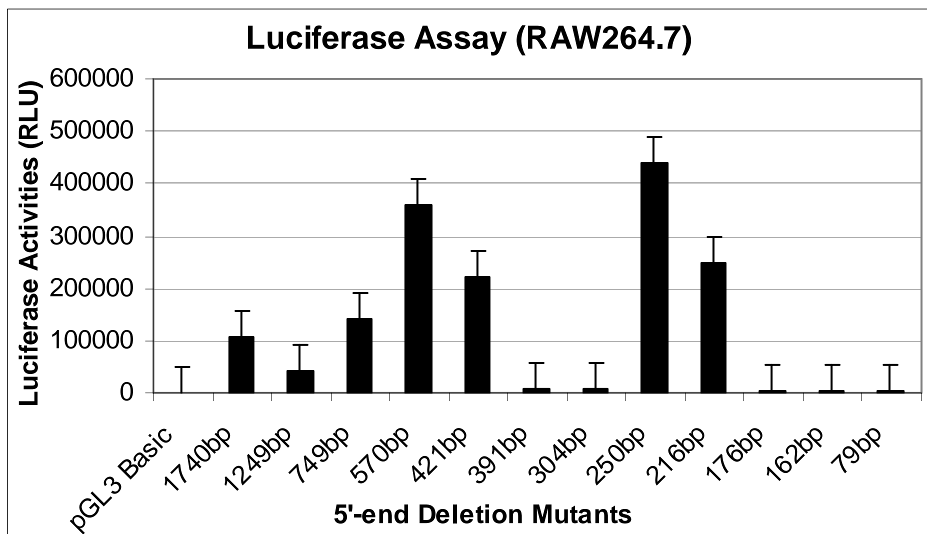

Fig 13C. Luciferase assay result of 5'-end deletion mutants in RAW264.7 cells.

Fig 13. Luciferase activity dropped in 5'-end deletion mutants shorter than 570bp indicating a possible enhancer between -453 and -274, while luciferase activity increased again in the 250bp and 216bp mutants suggesting a possible repressor located between fragment 304bp and 250bp. 5'-end deletion mutants shorter than 176bp lost luciferase activity, suggesting that the 216bp mutant may contain the core promoter of the SLA/LP gene.

To analyse the putative core promoter sequence, 3' end deletions of the 216bp mutant were performed. The luciferase assay results of the 3'-end deletion promoter fragments in three cell lines also showed a similar pattern of retained promoter activity of the 75bp and 63bp mutants and loss of promoter activity by the 54bp mutant, suggesting that the 9bp at 3'-

end of 63bp mutant was essential for transcription and the 63bp mutant had the basic function of initiating transcription (Fig 14).

However, there was a clear reduction of luciferase activity by the 75bp mutant, compared to the 216bp mutant, suggesting the presence of at least one activator in the 141bp region (-24 to +117) at the 3'-end of 216bp mutant.

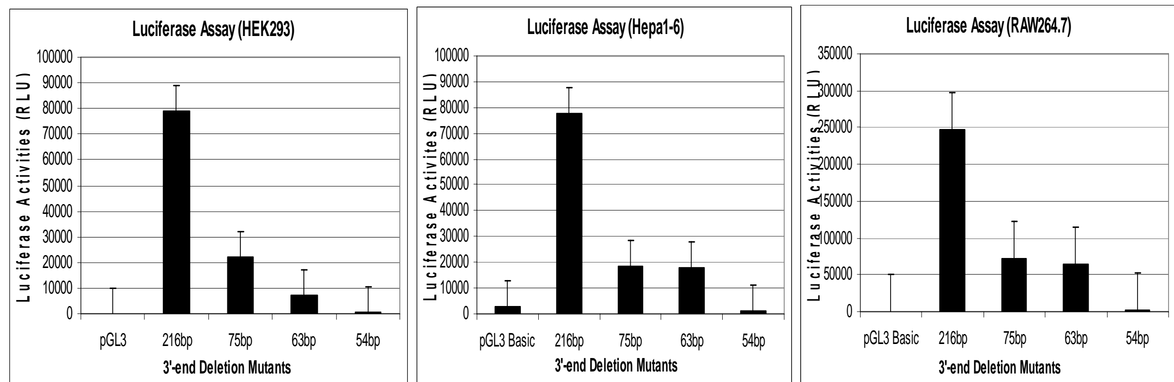

Fig 14. Luciferase assay results of 216bp, 75bp, 63bp and 54bp SLA/LP promoter mutants in HEK293, RAW264.7 and Hepa1-6 cells. Luciferase activity by the 75bp mutant was significantly decreased, compared to the 216bp mutant, indicating the presence of activator(s) within the -24 to +117 sequence. A 54bp mutant did not show luciferase activity, suggesting that the 63bp mutant contained the basic ability to initiate transcription.

To further characterise the putative core promoter, three 5'-deletion mutants of the 63bp mutant were produced (Fig 15).

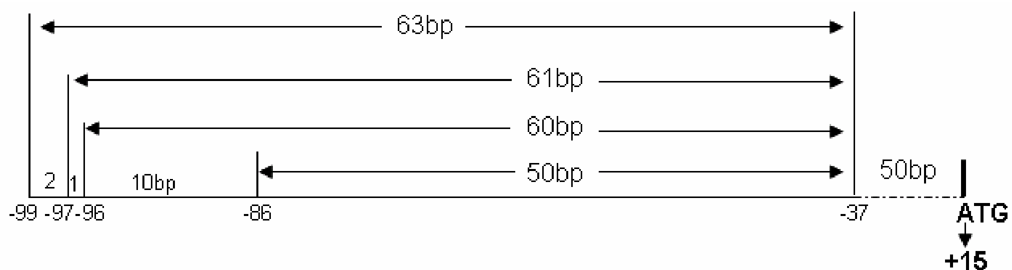

Fig 15. 5'-deletion of the 63bp SLA/LP gene fragment mutants generated by PCR. The length of the fragments was 63bp, 61bp, 60bp or 50bp.

Luciferase activity of the 50bp mutant was similar to the empty vector control. Luciferase activity of the 61bp and 60bp mutants were decreased, compared to the 63bp mutant (Fig 16). These results indicated that the first 13bp at the 5'-end of the 63bp mutant were critical for transcription, and that this ability was decreased when the first two or three base pairs was

deleted. This 63bp fragment seems to be the minimal sequence of SLA/LP gene that has the ability to initiate basic protein transcription.

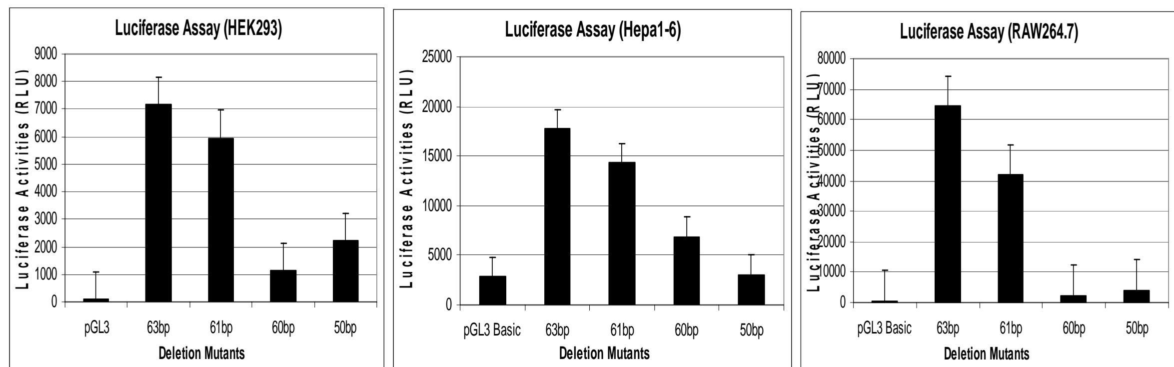

Fig 16. Luciferase assay results of the 63bp, 61bp, 60bp and 50bp mutants in HEK293, RAW264.7 and Hepa1-6 cells. Luciferase activity decreased when two or three base pairs at the 5'-end of the 63bp fragment were deleted, and significantly decreased when deleted 13bp at the 5'-end of the 63bp mutant. These results indicated that the 63bp fragment might be the minimal stretch to provoke transcription.

3.6 Predicted transcription factors and promoter type

The AliBaba 2.1 program was used to predict transcription factor binding sites with the 304bp fragment (-187 to +117). Within the 63bp fragment (-99 to -37), there were two Sp1, one RAP1 and one WT1 consensus sequences predicted. Whithin the region (-187 to -133) containing a possible repressor element, a TBP, an Oct-1, a C/EBPalp and a RAP1 binding sites were predicted. According to the luciferase assay results, Oct-1(-184 to -175), Sp1 (-85 to -76), RAP1 (-71 to -62) and Sp1 (-55 to -41) binding sites were tested with point mutations and gel shift assays.

The original prediction by the Alibaba2.1 program is shown in Fig17. The 1740bp SLA/LP promoter fragment and putative binding sites are shown in Fig 18.

AliBaba2.1 predicts the following sites in your sequence

```
============================== -187(304bp) =======================================
seq(    0..    59)     gccaactctctcatatatactcaaggtttccaacttcttcccaagtgcttgtggtgttgg
Segments:
4.5.1.0    10   19             ====TBP===
3.1.2.2    12   21               ===Oct-1==
1.1.3.0    19   28                      =C/EBPalp=
3.5.1.2    52   61                                                        ====RAP1
======================== -133(250bp) ===================== -99(216bp) ============
seq(   60..   119)     ggaatgggagtggagggccaccaagacggctctgcgcatgcgcggaagtcgaaggcggcg
Segments:
3.5.1.2    52   61     ==
2.3.1.0    69   78             ====Sp1===
9.9.637    92   101                                ===NRF-1==
```

Fig 17. The predicted transcription factors by AliBaba2.1 algorithm in the 304bp SLA/LP promoter fragment.

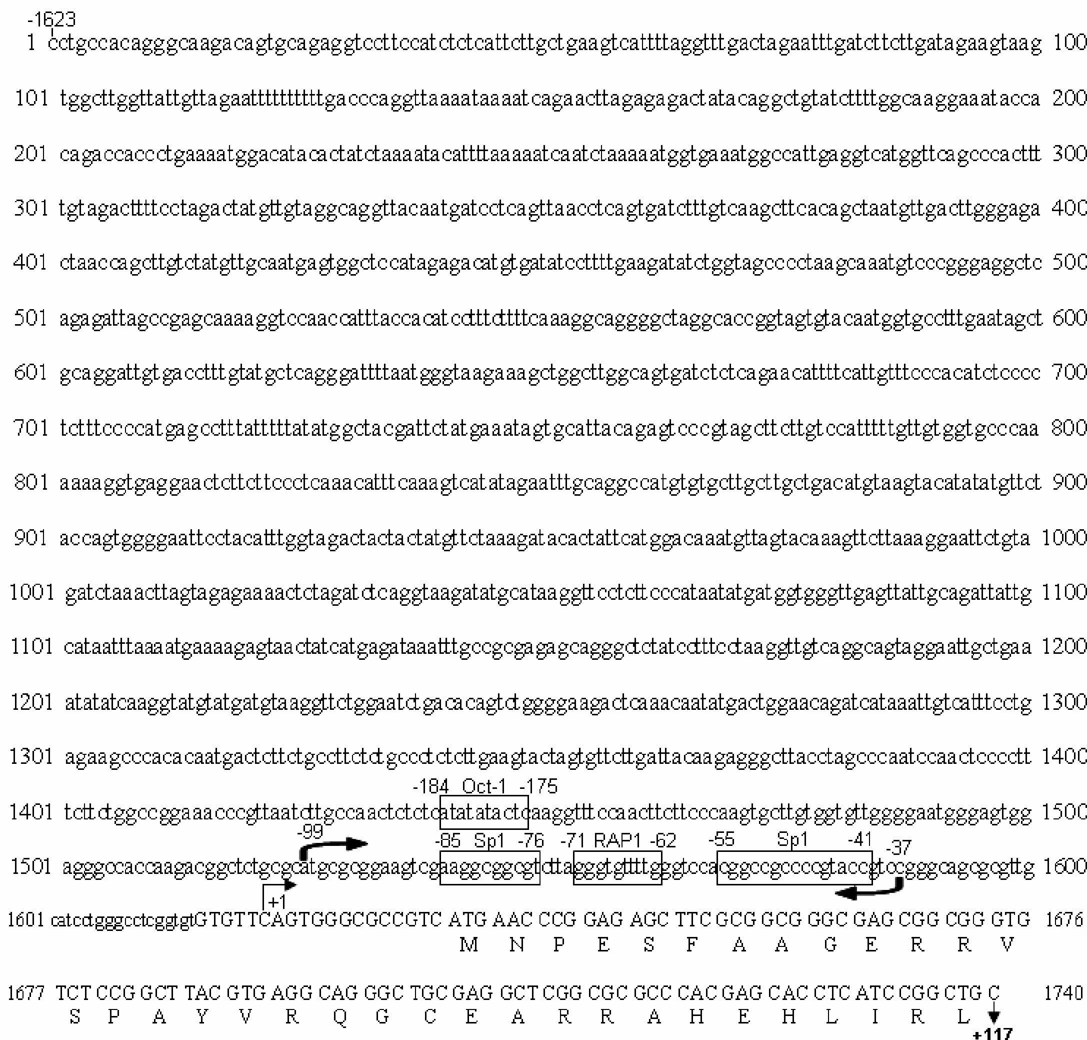

Fig 18. The 1740bp fragment used for promoter assay. Within the 63bp fragment (-99 to -37), there were two Sp1 (-85 to -76 and -55 to -41) and one RAP1 (-71 to -62) predicted. There was one Oct-1 binding site predicted at -184 to -175.

A CpG island promoter detection algorithm, CpGProD, indicated a high possibility for the SLA/LP promoter to be a CpG island promoter (Table 5). The CpG island starts from -508bp to +117bp of the SLA/LP gene. This CpG island length is 625bp, G+C frequency is 54.2%, observed / expected CpG ratio is 0.7843. The predicted possibility of promoter to be located in this CpG island was as high as 0.8051 (range from 0.5 to 1) and predicted probability to be located over the transcription start site was 0.8378 (range from 0 to 1).

Table 5 The Possibility to Be a CpG Island Promoter Predicted by CpGProD Program

sequence name	**number**	**begin**	**end**	**length (bp)**	**G+C frequency**	**CpGo/e ratio**	**start-p**	**AT skew**	**GC skew**	**strand (strand-p*)**
1740bp	1/1	1122	1746	625	0.5424	0.7843	0.8378	-0.0699	0.0383	plus (0.8051)

Strand (strand-p): strand of the promoter and predicted probability to be located over this strand (range from 0.5 to 1).

Start-p: predicted probability to be located over the transcription start site (range from 0 to 1).

3.7 Point mutations and luciferase assays

Different point mutations were introduced into the predicted transcription factor binding sites by PCR. Five single point mutantions, named Oct1-1Mu, Sp1-111Mu, Sp1-121Mu, RAP1-T Mu and Sp1-21Mu, were introduced into the putative transcription factor binding sites separately. One single point mutation at –89, named -89G to A, was introduced into the 63bp sequence in order to test whether the sequence outside the putative binding site was essential for transcription (Fig 19A).

To investigate the difference of one or two points mutations for promoter activity, a two-position mutation, named Oct1-2Mu, was introduced into the putative Oct1 binding site (Fig 19B).

All mutations were obtainded by PCR with forward primer containing Kpn I site and reverse primer containing Xho I site. The PCR products were digested by Kpn I/ Xho I and subcloned into Kpn I/Xho I digested pGL3 basic vector and used for luciferase assay.

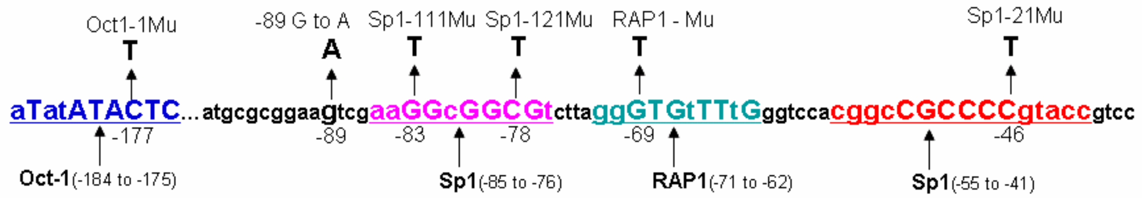

Fig 19A. Six clones with single point mutation at different positions of the 63bp fragment sequence. The clones were obtained by PCR with forward primer containing Kpn I site and reverse primer containing Xho I site. The PCR products were digested by Kpn I/ Xho I and subcloned into Kpn I/Xho I digested pGL3 basic vector.

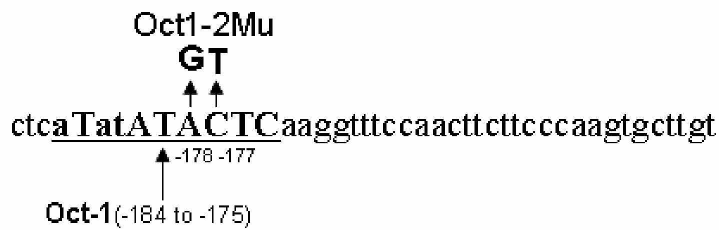

Fig 19B. Two-point mutation of the putative Oct-1 binding site. The mutation was obtained by PCR with forward primer containing Kpn I site and reverse primer containing Xho I site. The PCR products were digested by Kpn I/ Xho I and subcloned into Kpn I/Xho I digested pGL3 basic vector.

The luciferase assay of mutated putative Oct-1 binding site (-184 to -175) showed that luciferase activity was increased by single point mutation (Oct1-1Mu) 2.8-fold in HEK293cells, 3.5-fold in Hepa1-6 cells and 6.3-fold in RAW264.7 cells, compared to the 304bp mutant. Mutation at two positions (Oct1-2Mu) further increased luciferase activity by 5.5-fold, 4.6-fold and 6.8-fold in the three different cells (Fig 20A). Theses results indicate that the putative Oct-1 element serves as a repressor and may participate in SLA/LP transcription control.

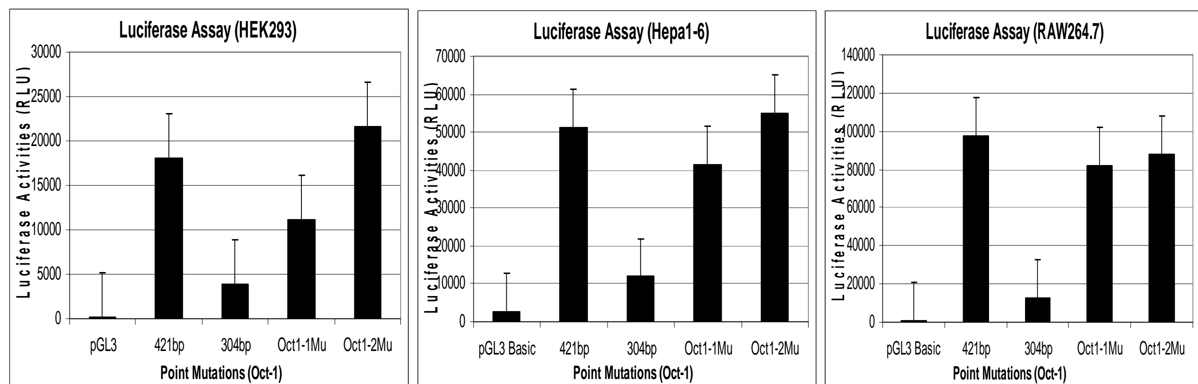

Fig 20A. Luciferase assay results of the Oct-1 mutations in HEK293, RAW264.7 and Hepa1-6 cells. Luciferase activity was significantly increased after mutation at a single position; additional mutation at a second position further increased luciferase activity. These results indicate that an Oct-1 repressor element may participate in SLA/LP transcription control.

The luciferase assay of mutated putative Sp1 binding site (Sp1-111Mu) showed that luciferase activity was decreased 9.1-fold in HEK293 cells, 4.4-fold in Hepa1-6 cells and 20.0-fold in RAW264.7 cells, compared to the 63bp native sequence. Another single point mutation, named Sp1-121Mu, yielded no decrease of luciferase activity in HEK293 cells, 2.9-fold lower in Hepa1-6 cells and 1.5-fold lower in RAW264.7 cells, compared to the 63bp native sequence (Fig 20B). The Sp1-111Mu mutation was at the 5'-end of the putative Sp1 binding site, while Sp1-121Mu mutation was at the 3'-end of this Sp1 bindind site. These results indicated that this putative Sp1-binding site is an enhancer sequence and the 5'-end of this Sp1 consensus sequence seems to be more critical for its function.

Luciferase activity vanished both after mutation of the putative RAP1 binding site at-71 to -62 (RAP1 T-Mu) or after mutation of the second putative Sp1 binding site at -55 to -41 (Sp1-21Mu), when compared to the 63bp native sequence, suggesting the existence of further enhancer elements (Fig 20B).

Luciferase activity was significantly decreased by single point mutation (-89G to A) outside the putative transcription factor binding sites, suggesting that the complete 63bp sequence (-99 to -37) seems to be necessary for transcription and the presence of another enhancer element not detected by the used search algorithms (Fig 20B).

These findings indicate that the transcription of SLA/LP gene may be regulated at least by the transcription factors Sp1, Oct-1 and RAP1.

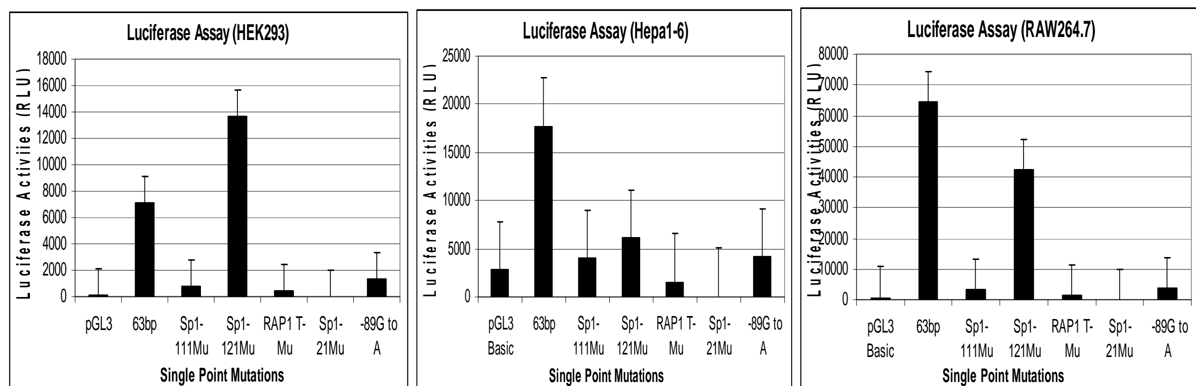

Fig 20B. Luciferase assay results of five single point mutations, Sp1-111Mu, Sp1-121Mu, RAP1 T-Mu, Sp1-21Mu and -89G to A in HEK293, RAW264.7 and Hepa1-6 cells. Luciferase activity decreased or vanished in mutant clone of RAP1(-71 to -62), Sp1 site at -55 to -41 and the 5'-end mutation of the Sp1 site at -85 to -76. These results indicate that these transcription factors may regulate transcription of SLA/LP gene.

3.8 Gel Shift Assays and Super Gel Shift Assays

Gel shift assays were conducted to further confirm the identified transcription factor binding sites; biotin-labeled oligonucleotides were used as probes.

Recombinant Sp1 protein caused a band shift of the -85 to -76 oligonucleotide sequence and an antibody to Sp1 induced a supershift of the band. The specific competitor of Sp1 consensus sequence (unlabeled Sp1 probe) decreased the binding ability of Biotin-labeled Sp1 probe. Thus, the -85 to -76 oligonucleotide sequence seems to bear an Sp1 element.

The second putative Sp1 binding site (-55 to -41) produced a weaker band shift when incubated with recombinant Sp1 protein, and no supershift with anti-Sp1 antibody. Nevertheless, the unlabelled competitor probe inhibited the band shift produced by the putative Sp1 binding site. Therefore, the -55 to -41 oligonucleotide probably bears a weak Sp1 binding element. See Fig 21A and Fig 21B.

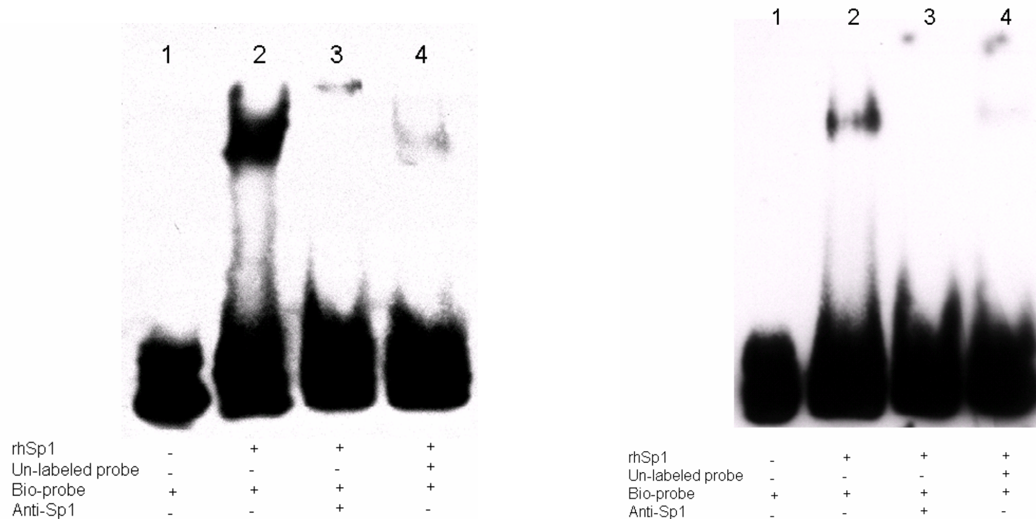

Fig 21A. Sp1 (-85 to -76) Gel Shift Assay

Fig 21B. Sp1 (-55 to -41) Gel Shift Assay

Fig 21. Gel Shift Assay and Super Gel Shift Assay of putative Sp1 consensus sequence. There was a band shift (lane 2) and the specific competitor decreased the band shift (lane 4) in Fig 21A and 21B. A super band shift showed (lane3) in Fig 21A but not in Fig 21B, indicating that the consensus sequence at position -85 to -76 may be a Sp1 factor binding site; while at position -55 to -41 might be a Sp1 binding site.

Lane1: negative control containing Biotin-labeled probe.

Lane2: Shift reaction containing Biotin-labeled probe and rhSp1 protein.

Lane3: Super shift reaction containing Biotin-labeled probe, rhSp1 protein and anti-Sp1 antibody.

Lane4: Specific competitor reaction containing Biotin-labeled probe, rhSp1 protein and unlabeled Sp1 probe.

Since recombinant Oct-1 was not available, we used HeLa nuclear extract to test for the presence of an Oct-1 element. HeLa nuclear extract induced a band shift of the putative Oct-1 binding site at position -184 to -175; this could be supershifted by Oct-1 specific antibody. The specific competitor oligonucleotide probe decreased the band shift. Thus, we may conclude that there is an Oct-1 binding site in the SLA/LP promoter region at position -184 to -175 (Fig 22).

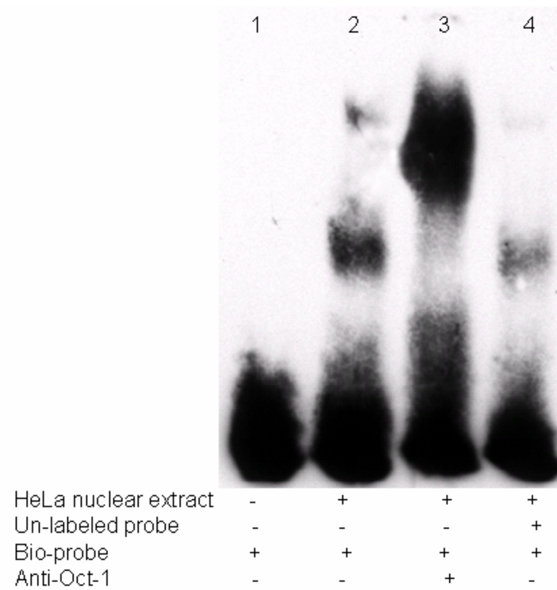

Fig 22. Gel Shift Assay and Super Gel Shift Assay of Oct-1. There was a shifted band in lane 2, a super shifted band in lane 3. The specific competitor decreased the band shift (Lane 4). These results indicated that this is an Oct-1 binding site.

Lane1: negative control containing Biotin-labeled probe.

Lane2: Shift reaction containing Biotin-labeled probe and HeLa nuclear extract.

Lane3: Super shift reaction containing Biotin-labeled probe, HeLa nuclear extract and anti-Oct-1 Antibody

Lane4: Specific competitor reaction containing Biotin-labeled probe, HeLa nuclear extract and unlabeled Oct-1 probe

To test for a putative RAP1 binding site, we also used HeLa nuclear extract. HeLa nuclear extract produced four shifted bands with the putative RAP1 binding sequence, which probably indicates the binding of several DNA-binding proteins contained in the HeLa nuclear extract. The specific competitor could inhibit at least one of these shifted bands. There was no super shift with Anti-RAP1 antibody (Fig 23). However, the only available

RAP1 antibody has not been tested for function in band shift assays. Therefore, it is likely but not certain that the -71 to -62 oligonucleotide sequence contained a RAP1 binding site.

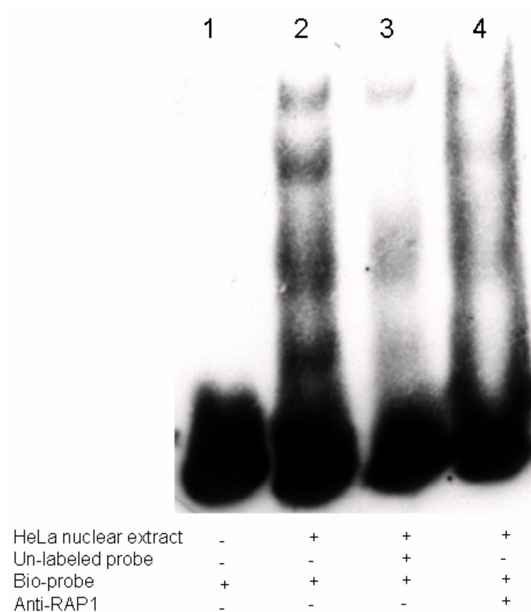

Fig 23. Gel Shift Assay and Super Gel Shift Assay of RAP1.

Lane1: negative control containing Biotin-labeled probe.

Lane2: Shift reaction containing Biotin-labeled probe and HeLa nuclear extract.

Lane3: Specific competitor reaction containing Biotin-labeled probe, HeLa nuclear extract and unlabeled RAP1 probe

Lane4: Super shift reaction containing Biotin-labeled probe, HeLa nuclear extract and anti-RAP1 Antibody

There were four shifted bands after incubated with HeLa nuclear extract (lane 2), which could be inhibited by specific competitor (lane 3). There was no super shift with Anti-RAP1 antibody which was not tested for function in gel shift assay (lane 4). The results indicated that this binding factor likely but not certain to be RAP1.

4. Discussion

The question of how a gene is expressed differentially is the key to our understanding of genetic regulation. There are two broad classes of genetic elements that control transcription. One class is the cis-elements; the other class is the trans-elements.

The cis-regulatory elements, which are DNA sequences in the vicinity of the structural portion of a gene that are required for gene expression, include the following: the core promoter, which recruits the transcription machinery and directs accurate initiation of transcription; enhancer or silencer sequences, which activate or inhibit transcription in different tissues; locus control regions, originally functionally defined as dominant activating sequences that confer position-independent and copy number-dependent expression on a linked transgene in transgenic mice (Grosveld et al. 1987); and insulators/boundary regions, proposed to be required at the borders of a regulatory domain of a gene to counteract inappropriate effects of nearby heterochromatin and/or distal enhancers (West et al. 2002; Labrador et al. 2002).

The trans-elements, which are usually considered being proteins that bind to the cis-acting sequences to control gene expression, include the following: basal transcription factors, TFII-A, B, D, E, F, H, which bind RNA Polymerase to stabilize the initiation complex; mediator, RNA pol II holoenzyme, which form the initiation complex; sequence specific transcription factors- activators and repressors, which bind to promoters at specific sequences (Carey 1998; Smale et al. 1998).

Activators are generally modular proteins containing a single DNA-binding domain and one or a few activation domains. Most of the DNA-binding domains have characteristic consensus amino acid sequences. An activation domain is a polypeptide sequence that activates transcription when it is fused to a DNA-binding domain. The different domains frequently are linked through flexible polypeptide regions. This may allow activation domains in different activators to interact even when their DNA-binding domains are bound to sites separated by tens of base pairs. Most eukaryotic repressors also are modular proteins. Similar to activators, they usually contain a single DNA- binding domain, one or a few repression domains, and can control transcription when they are bound at sites hundreds to thousands of base pairs from a start site (Lodish et al. 2000).

Transcription factors often are classified according to the type of DNA-binding domain. There are mainly homeodomain proteins, bZIP (basic zipper) proteins, bHLH (basic helix-loop-helix) proteins, and several types of Zinc finger proteins according to the most common structural motifs. Three types of DNA-binding proteins can form heterodimers: C4 zinc-

finger proteins, bZIP proteins, and bHLH proteins. The ability of these transcription factors to form heterodimers increases the number of DNA sites from which these factors can control transcription and the ways they can be controlled (Lodish et al. 2000).

Activators are more common and well studied than repressors. Both activators and repressors are important in regulating transcription. The absence of appropriate repressor activity can have devastating consequences. For instance, the protein encoded by the Wilms' tumor (WT1) gene is a repressor that is expressed preferentially in the developing kidney. Children, who inherit mutations in both the maternal and paternal WT1 genes, produce no functional WT1 protein; they invariably develop kidney tumors early in life (Lodish et al. 2000).

This study aimed to map the functional promoter sequence domains of the murine SLA/LP gene and determine which sequences are bound by presumable trans-acting factors during the initiation of expression.

A 1740bp fragment (-1623bp to +117bp) of mouse SLA/LP gene was cloned into upstream of luciferase in the reporter vector – pGL3 Basic vector. The high luciferase activity of this clone showed us clearly that the SLA/LP promoter was located in this 1740bp fragment and was orientation dependent. However, this 1740bp fragment does not contain the typical core promoter motifs, such as TATA box, DPE, BRE and Inr, at appropriate locations upstream of SLA/LP transcription start site. According to a CpG island standard set by Gardiner-Garden and Frommer, length at least 200bp, G+C content greater than 50% and the observed/expected CpG ratio greater than 0.6 (Gardiner-Garden et al. 1987), there is a CpG island starting from -508bp to +117bp of the SLA/LP gene. The predicted possibility of the promoter to be located in this CpG island was as high as 0.8051 (range from 0.5 to 1) and predicted probability to be located over the transcription start site was 0.8378 (range from 0 to 1) by a CpG island promoter detection algorithm-CpGProD. The transcription start site is actually located in this CpG island according to gene bank (Gene ID: 211006).

According to the hypothesis that CpG islands are associated with promoters that are transcriptionally active at totipotent stages of development (Antequera et al. 1999), we further believed that SLA/LP promoter was a CpG island promoter because we detected the mouse SLA/LP mRNA as early as 7th embryo day with Northern blot analysis.

To identify sequences important for promoter activity, a series of 5'-end deletion constructs were made by digestion with restriction enzymes or by PCR to synthesize the serial 5'-end deletion mutant fragments of 1249bp (-1132 to +117), 749bp (-632 to +117), 570bp (-453 to +117), 421bp (-304 to +117), 391bp (-274 to +117), 304bp (-187 to +117),

250bp (-132 to +117), 216bp (-99 to +117), 176bp (-59 to +117), 162bp (-45 to +117) and 79bp (+39 to +117) based on the 1740bp (-1623 to +117) fragment. The luciferase activities of these clones showed a similar pattern in three cell lines.

The relevant luciferase activity dropped with 391bp mutant, and further dropped with 304bp mutant; while increased again from 250bp mutant. This phenomenon indicated that there may be a repressor sequence located between 304bp and 250bp mutant which is 54bp in length. The transcription prediction algorithm, AliBaba2.1 predicts an Oct-1 binding site within this region. The luciferase activities decreased dramatically from 216bp mutant to 176bp mutant, suggesting that the 40bp at the 5'-end of 216bp mutant was a critically important fragment. Then, this 40bp was subcloned into pGL3 Basic vector and it did not show luciferase activity (data not shown), meaning it was not enough to initiate transcription and the fragment should be longer than 40bp in order to provoke transcription.

The 216bp mutant was used as a template to generate further 3'-deletion by PCR, resulting in 75bp (-99 to -25), 63bp (-99 to -37) or 54bp (-99 to -46) mutants. The luciferase assays gave us very interesting results. First, there was no big difference of the luciferase activities between 75bp mutant and 63bp mutant, while the 54bp mutant lost the activity completely. It can be deduced that the 9bp at 3'-end of the 63bp mutant were very important and the 63bp mutant might be the minimal fragment sequence starting transcription. Second, luciferase activity of the 216bp mutant was much higher than of the 75bp and 63bp mutants which might mean that there was at least one enhancer located at the last 141bp (-24 to +117) of 3'-end of 216bp mutant. Moreover, AliBaba2.1 predicted several Sp1 binding sites downstream of the transcription start site. We did not test the existence of these putative Sp1 factors in our experiment.

The 63bp mutant, which could start transcription of luciferase, was used to get further 5'-deletion fragments of 61bp (-97 to -37), 60bp (-96 to -37) or 50bp (-86 to -37). The luciferase assay results of these clones showed that the luciferase activity decreased with clone 61bp and 60bp, while 50bp lost the ability. Moreover, the single point mutation at -89 (G toA) significantly decreased luciferase activity. Thus, it was confirmed again that this 63bp mutant was the minimal stretch sequence to initiate luciferase transcription in vitro and it might be the core promoter region of SLA/LP gene.

The 63bp sequence, from nucleotide -99 to -37, contains two putative binding sites for transcription factor Sp1 (nucleotides -85 to -76 and -55 to -41) and one RAP1 binding site (-71 to -62) based on the prediction results of transcription prediction algorithm, AliBaba2.1. Single point mutation (purine changed into thymine) was introduced into the 63bp sequence

by PCR at the predicted binding sites. Mutation of the Sp1 (-55 to -41) or RAP1 (-71 to -62) sequences abrogated the ability to initiate luciferase transcription. For the Sp1 binding site at -85 to -76, the mutation could not initiate the expression effectively when the mutation was introduced into the 5'-end of the Sp1 binding site; however, there was no obvious change of luciferase activity when the mutation was at the 3'-end of this Sp1 binding site. This suggested that the 5'-end of this Sp1 site is more critical than 3'-end to maintain its function. In addition, a single base mutation at -89 (G → A) of this 63bp sequence also impaired the luciferase expression, indicating that an additional, so far not defined element may be involved in the promoter function. Together, these findings indicate at least a role for Sp1 and RAP1 in activating transcription of SLA/LP gene.

The Oct-1 binding site (-184 to -175) was located 75bp upstream of the 63bp presumable core promoter sequence. A single point mutation could increase the transcription activity by 2.8 to 6.3-fold in the different cell lines, while two point mutations further increased transcription ability by 4.6 to 6.8-fold. These results indicate that Oct-1 may be a repressor of SLA/LP gene expression.

All luciferase assays were performed with co-transfection of ß-galactosidase into three different cell lines and repeated at least three times. The changes of relevant luciferase activity had the similar patterns in three cell lines, human embryonic kidney epithelial cells, mouse hepatocellular cells and murine macrophage cells. We may thus conclude that the SLA/LP promoter activity had no cell type specificity, which could be one of the reasons why SLA/LP is expressed widely in different tissues in vivo.

The gel shift assay of the Sp1 binding site at -85 to -76 confirmed the results of the luciferase assays droving by the different mutants. The Biotin-labelled Sp1 probe (-96 to -72) gave a strong band shift after incubation with recombinant human Sp1 protein. This shifted band could be specifically inhibited by competition with an unlabeled probe. The presence of Sp1 protein in this complex was further confirmed by incubation with polyclonal Anti-Sp1 antibodies (rabbit anti-human). The Sp1-shifted band was supershifted by Anti-Sp1 antibodies. These findings indicated that the Sp1 protein bound specifically at the Sp1 binding site (-85 to -76) in vitro and thus confirmed the existence of a Sp1 element in the SLA/LP promoter region.

Recombinant Sp1 protein also caused a band shift of the -55 to -41 oligonucleotide sequence, but no supershift was observed with anti-Sp1 antibody. Nevertheless, the unlabelled competitor probe inhibited the band shift produced by the putative Sp1 binding site (-55 to -41). Therefore, the -55 to -41 oligonucleotide likely, but not certainly bears a

Sp1 binding element. This Sp1-like element played a critical role in SLA/LP promoter activity because single base mutation abolished the promoter function totally.

Sp1 is ubiquitously expressed in mammalian cells and participates in regulating the expression of genes involved in almost all cellular processes (Cawley et al. 2004). Sp1 knockout mice revealed that Sp1 is essential for normal mouse embryogenesis. All $Sp1^{-/-}$ mice die at approximately day 11 of gestation (Marin et al. 1997). The presence of Sp1 binding sites in CpG islands is particularly notable. Not only does Sp1 contribute to the maintenance of the hypomethylated start of CpG islands (Brandeis et al. 1994; Macleod et al. 1994), but it may also function in concert with the basal transcription factors to mediate transcription initiation. Transcription start sites are often located 40-80bp downstream of the Sp1 sites; this suggests that Sp1 may direct the basal machinery to form a preinitiation complex within a loosely defined window (Smale et al. 1990; Blake et al. 1990). The presence of probable Sp1 elements at position -55 to -41 and -85 to -76 is compatible with this notion. Sp1 controlled genes mostly code general metabolic proteins and not cell-specific proteins. It is consistent with the expression pattern of SLA/LP protein.

The molecular properties of Sp1 have been studied in vitro in detail. The protein is phosphorylated (Jackson et al. 1990) and highly glycosylated (Jackson et al. 1988). The N-terminus contains glutamine-and serine/threonine-rich domains that are essential for transcriptional activation (Courey et al. 1989). The C-terminal domain of Sp1 is involved in synergistic activation and interaction with other transcription factors (Li et al. 1991). Sp1 may be required for the maintenance of differentiated cells as $Sp1^{-/-}$ ES cells can contribute extensively to every tissue of E9.5 (embryonic day 9.5) chimeras but not to later embryonic time points or newborn mice (Marin et al. 1997).

The absence of a TATA box from the SLA/LP promoter and the likely presence of CpG island are consistent with the finding that Sp1-binding sites are often located close to the transcription initiation site in TATA-less promoters and contribute to transcriptional initiation (Butler et al. 2002).

HeLa nuclear extract produced four shifted bands with the putative RAP1 binding sequence, which was probably due to binding of several DNA-binding proteins contained in the HeLa nuclear extract. The specific competitor could inhibit at least one of these shifted bands by the unlabeled competent probe. There was no super shift with Anti-RAP1 antibody (mouse anti-human monoclonal antibody). However, the only available RAP1 antibody has not been tested for function in band shift assays. Therefore, it is likely but not certain that the -71 to -62 oligonucleotide sequence contained a RAP1 binding site. The other possible

explanation might be that this RAP1 factor indirectly bound other transcription factors binding the same oligonucleotide-probe and this protein-protein interaction caused the DNA binding was completely abolished by the antibodies. At least, the -71 to -62 ologonucleotide sequence was a cis-element, which might contain a RAP1 binding site or that of another unknown factor. Single point mutation of this site dramatically decreased luciferase activity, indicating that it is an important activator-binding site.

The RAP1 protein was initially identified as a factor that binds to the HMRE silencer element and to the MATalpha UAS (Shore et al. 1987, 1987; Buchman et al. 1988). This same RAP1 binding site from the HMR locus is, however, able to activate transcription from heterologous promoters (Brand et al. 1987). The function of RAP1 in transcription regulation may vary. The specific sequence of a RAP1-binding site does not determine its function (Brand et al. 1987; Shore et al. 1987; Buchman et al. 1988).The function of activating and repressing transcription depends on the context in which its DNA-recognition sequence is placed, via interactions with nearby DNA-binding proteins and other specific trans-acting regulators (Shore 1994). The mechanism for transcriptional silencing by RAP1 depends on its interaction with a set of silencing proteins, Sir2p, Sir3p and Sir4p. It is limited to specific regions of the chromosomes and apparently, it only occurs in certain compartments within the yeast nucleus (Palladino et al. 1993). In contrast, transcriptional activation by RAP1 is much more widespread, involving many genes that encode, inter alia, glycolytic enzymes and ribosomal proteins (Woudt et al. 1987; Nieuwint et al. 1989; Tornow et al. 1990; Shore 1994). However, at some promoters, RAP1 doesn't act directly as a transcriptional activator, but instead acts as a factor that allows binding by other regulatory proteins (Morse 2000). For example, RAP1 and another DNA-binding protein, Gcr1p, act synergistically to activate transcription of many glycolytic enzyme genes (Baker 1986, 1991; Tornow et al. 1990). The knowledge of how RAP1 regulates this large number of genes, many of them with different expression patterns, is very helpful to understanding the transcriptional regulation patterns.

According to our results, the putative trans-acting factor, RAP1, is an activator for initiating SLA/LP transcription. Moreover, it might act synergistically with Sp1 to activate transcription of SLA/LP gene, since mutations at each binding site alone could lead to the loss of promoter activity. It would be interesting to know how RAP1 co-acted with Sp1 trans-acting element in this process.

Oct-1 interacts with an 8bp sequence termed the octamer motif (5'-ATGCAAAT-3') and related sequences (Verrijzer et al. 1992) but with different affinity. In our study, the

sequence of the SLA/LP gene contains an imperfect consensus octamer-binding site, aTatATACTC, in the promoter region. However, we could show that Oct-1 transcription factor can interact with this DNA sequence by gel shift assay and super gel shift assay. Thus, Oct-1 binding to this sequence may act as a silencer in the SLA/LP promoter. Indeed, single point mutation increased the promoter activity by 2.8 to 6.3-fold in different cell lines, and two point mutations further increased this.

Oct-1 protein is ubiquitously expressed and involved in the regulation of various genes, which function in the development of multiple organs and tissues (Roberts et al. 1991; Fadel et al. 1999), the control of cell cycle progression (Roberts et al. 1991), and the regulation of signalling pathways as well (Prefontaine et al. 1998; Kakizawa et al. 1999). It has been shown that Oct-1 acts not only as a transcriptional activator (Kim et al. 1995; Oren et al. 2005) but also as a transcriptional repressor for certain genes. For example, VCAM-1 (Iademarco et al. 1993), von Willebrand factor promoter (Schwachtgen et al. 1998), pro-lactin gene promoter (Subramaniam et al. 1998), rGH promoter (Kakizawa et al. 1999), interleukin-8 expression (Wu et al. 1997; Zhang et al. 1999), virus-induced Interferon alpha gene expression (Mesplède et al. 2005) and the collagenase gene which is one of the cellular aging-associated genes (Imai et al. 1997) are shown to be down-regulated by Oct-1. In this study, we identified Oct-1 as a repressor of SLA/LP gene expression. However, the mechanism of the bifunctional transcriptional activity of Oct-1 was not fully understood. In certain binding conformations, Oct-1 factor may cooperate with activating co-regulators, while Oct-1 can also recruit inhibitory molecules. OCAB/OBF-1 has been shown to be involved in the transcriptional activation by Oct-1, whereas the silencing mediator for retinoid and thyroid hormone receptors (SMRT) interacts with Oct-1 and is involved in transcriptional repression by Oct-1 (Kakizawa et al. 2001).

It is reported that Oct-1 and Sp1 may physically interact and cooperatively stimulate expression in the case of the small nuclear U2 RNA gene (Janson et al. 1990; Strom et al. 1996). However, in our case, we presumed that Oct-1 might interact with a silencing mediator located outside the SLA/LP core promoter region to down regulate the expression of SLA/LP gene. Therefore, we may presume that the normal expression of SLA/LP is a result of the balance between the activator Sp1/RAP1 and repressor Oct-1.

Transcription factors can directly regulate assembly of transcription-initiation complexes and the rate at which they initiate transcription. The concentrations and activities of activators and repressors that control transcription of many protein-coding genes are regulated during cellular differentiation and in response to hormones and signals from

neighbouring cells (Lodish et al. 2000). These three levels of control by transcription factors regulate the frequency of transcription initiation for the genes that are transcribed in different cell types. Therefore, it would be very helpful to understand the high expression of SLA/LP protein in liver/pancreas, if we further test the concentration and activity of Sp1, RAP1 and Oct-1 in these organs, or find some liver/pancreas specific signals regulating these transcription factor activities.

We did not find a cell type or tissue specific transcription factors in this study as Sp1, RAP1 and Oct-1 are expressed ubiquitously. This could be the reason why the SLA/LP protein is expressed in many different tissues.

The promoter architectures can be divided into four types which provide clues to regulatory mechanisms according to a study of yeast genome (Harbison et al. 2004). The first type of a promoter structure is single regulator architecture which is often involved in a common biological function. The second type is repetitive motifs which are necessary for stable binding and can permit a graded transcriptional response. The third type is multiple regulator architecture which implies that the gene might be subject to combinational regulation. The fourth type is co-occurring regulator architecture, promoters that contain binding site sequences for recurrent pairs of regulators, which implies that the two regulators interact physically or have related functions at multiple genes. The structure of SLA/LP promoter is more like a multiple regulators architecture in which at least three or four transcription factors are important to regulate the expression of SLA/LP protein. At the same time, it also means that SLA/LP gene is regulated in a combinational way.

Lack of classical TATA or CCAAT boxes, an increased GC content with functional Sp1 site(s) in the proximal promoter region, and a CpG island close to the transcriptional initiation site are promoter features typical of housekeeping genes such as glycolytic enzymes, thymidylate synthase, adenine deaminase, and dihydrofolate reductase (Swick et al. 1989; Rundlof et al. 2001), the latter being studied as a model housekeeping promoter (Jensen et al. 1997). In addition, Oct-1 is ubiquitously expressed and believed to govern the transcription of many housekeeping genes (Ryan et al. 1997). Housekeeping genes are turned on early in fetal development and stay on throughout adulthood in almost all tissues (Warrington et al. 2000; Zhang et al. 2004). The expression level of housekeeping gene has been shown significant differences between tissue types and between donors of the same tissue (Barber et al. 2005). SLA/LP mRNA is also well shown to be expressed in many diverse tissues and early embryos by Northern blotting. The transcriptional activity and structure of the core promoter of SLA/LP analyzed in the present study fulfil house keeping

gene features. Based upon this expression pattern in combination with the structure and function of the core promoter as described herein, we thereby propose that SLA/LP gene should be considered to belong to the class of housekeeping genes that maintain essential cellular functions.

In general, we have defined SLA/LP gene promoter region and several transcription factors in this study. The SLA/LP promoter region lacks the typical core promoter elements, such as TATA box, Inr, BRE and DPE. The minimal promoter region, or the core promoter, is 63bp at -99 to -37 in the 5' flanking region of transcription start site in SLA/LP gene. This core promoter is required for basal promoter activity in cells. Sequence analysis of this 63bp region showed a GC-rich region with one consensus binding site for Sp1, one Sp1-like factor binding site and a potential RAP1 transcription factor binding site. All of these three cis-activating elements are necessary for full promoter activity in vitro as the point mutation abolished the initiation of transcription. According to our study, we may conclude that Sp1 functions synergistically with another transcriptional factor, putatively RAP1, to start the transcription of SLA/LP gene. There is another cis-element, Oct-1 binding site, 75bp upstream of this core promoter region positioned at -184 to -175. This cis-element works as a repressor and the point mutation could increase the transcription significantly. This Oct-1 transcription factor might be a key factor down-regulating the expression of SLA/LP protein. The SLA/LP promoter is probably a CpG island promoter which usually contains multiple transcription start sites. It has a multiple regulator architecture, which indicates the complicated regulation pattern.

The observation that SLA/LP mRNA is transcribed in early embryo indicates it plays important roles during embryonic development. The non-tissue specific expression of SLA/LP protein is consistent with the finding that the transcription factors found in the SLA/LP core promoter are not tissue specific, which may indicate a more general role of SLA/LP protein in different tissues. The expression of SLA/LP is regulated by different transcription factors, Sp1 and RAP1 being probable activators, and Oct-1 being a likely repressor. The high expression of SLA/LP protein in liver/pancreas as well as the reported up-regulation in activated lymphocytes might be caused by high activity and concentration of the activators, weak control of repressor and/or some liver/pancreas specific signals.

According to the expression pattern and the features of SLA/LP promoter, it implicates strongly that SLA/LP gene may be a housekeeping gene.

It remains unclear how the loss of immune tolerance to SLA/LP in autoimmune hepatitis is initiated. It is currently believed that this is largely at the level of peripheral tolerance and

may involve abnormalities of CD25 positive T suppressor cells, and inappropriate expression of co-stimulatory signals (Burt et al. 2004). Thus, it would be helpful to understand the pathogenesis and autoimmunity of AIH, if we can compare the expression of these transcription factors in liver biopsies from both normal tissues and AIH patients.

5. Summary

Anti-SLA/LP antibodies are thus far the only known autoantibody, which are specific markers of autoimmune hepatitis (AIH) with 100% specificity and 20% sensitivity (Wies et al. 2000). However, the function, regulation of expression and the pathogenetic role of SLA/LP are not clear. It will be helpful for understanding autoimmunity to SLA/LP and the pathogenesis of AIH, if we knew more about SLA/LP. The SLA/LP protein is expressed widely in different normal tissues and mouse SLA/LP mRNA can be detected as early as day 7 of embryonic development. Therefore, it seems to have a role during embryogenesis and in the function of adult tissues.

In this study, we focused on the transcriptional regulation of murine SLA/LP protein. We mapped the core promoter region of the SLA/LP gene, which is located at -99 to -37 flanking the 5'-end of transcription start site. Within this core promoter, a Sp1 binding site, a Sp1-like binding site and a possible RAP1 binding site was identified. All of these cis-activating elements are essential for maintaining the basic initiation of transcription of SLA/LP protein. 75bp upstream of this core promoter, there is an Oct-1 binding site that acts as a repressor down-regulating the expression of SLA/LP protein.

SLA/LP gene promoter is a CpG island promoter, which normally contains multiple transcription start sites. This promoter lacks typical core promoter elements, such as TATA box, Inr, DPE and BRE. The structure of the SLA/LP promoter is a multiple regulator architecture that means that the SLA/LP gene is regulated in a complicated combinational way.

The expression of SLA/LP protein seems to be the result of balance between activator Sp1/RAP1 and repressor Oct-1.

SLA/LP has also been confirmed to be an important protein during embryonic development. The non-tissue specific expression of SLA/LP protein may be due to the non-tissue specific transcription factors. The concentration and activity of these factors as well as some liver/pancreas specific signals may contribute to the translation frequency of SLA/LP protein in different tissues.

The expression pattern and the promoter features suggest strongly that the mouse SLA/LP gene may be a housekeeping gene.

6. References

Advani SJ, Durand LO, Weichselbaum RR, Roizman B (2003) Oct-1 is posttranslationally modified and exhibits reduced capacity to bind cognate sites at late times after infection with herpes simplex virus 1. J Virol 77:11927-11932

Alvarez F, Berg PA, Bianchi FB, Bianchi L, Burroughs AK, et al (1999) International autoimmune hepatitis group report: review of criteria for diagnosis of autoimmune hepatitis. J Hepatol 31:929-938

Antequera F and Bird A (1993) Number of CpG islands and genes in human and mouse. Proc Natl Acad Sci USA 90:11995-11999

Antequera F and Bird A (1999) CpG islands as genomic footprints of promoters that are associated with replication origins. Curr Biol 9:R661-667

Arkhipova IR (1995) Promoter elements in Drosophila melanogaster revealed by sequence analysis. Genetics 139:1359-1369

Baeres M, Herkel J, Czaja AJ, Wies I, Kanzler S, Cancado EL, Porta G, Nishioka M, Simon T, Daehnrich C, Schlumberger W, Galle PR, Lohse AW (2002) Establishment of standardised SLA/LP immunoassays: specificity for autoimmune hepatitis, worldwide occurrence, and clinical characteristics. Gut 51:259-264

Baker HV (1986) Glycolytic gene expression in Saccharomyces cerevisiae: nucleotide sequence of GCR1, null mutants, and evidence for expression. Mol Cell Biol 6:3774-3784

Baker HV (1991) GCR1 of Saccharomyces cerevisiae encodes a DNA binding protein whose binding is abolished by mutations in the CTTCC sequence motif. Proc Natl Acad Sci U S A 88:9443-9447

Barber RD, Harmer DW, Coleman RA, Clark BJ (2005) GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. Physiol Genomics 21:389-395

Bellorini M, Dantonel JC, Yoon JB, Roeder RG, Tora L, Mantovani R (1996) The major histocompatibility complex class II Ea promoter requires TFIID binding to an initiator sequence. Mol Cell Biol 16:503-512

Bird AP (1986) CpG-rich islands and the function of DNA methylation. Nature 321:209-213

Black AR, Jensen D, Lin SY, Azizkhan JC (1999) Growth/cell cycle regulation of Sp1 phosphorylation. J Biol Chem 274:1207-1215

Blake MC, Jambou RC, Swick AG, Kahn JW, Azizkhan JC (1990) Transcriptional initiation is controlled by upstream GC-box interactions in a TATAA-less promoter. Mol Cell Biol 10:6632-6641

Bouwman P and Philipsen S (2002) Regulation of the activity of Sp1-related transcription factors. Mol Cell Endocrinol 195:27-38

Brand AH, Micklem G, Nasmyth K (1987) A yeast silencer contains sequences that can promote autonomous plasmid replication and transcriptional activation. Cell 51:709-719

Brandeis M, Frank D, Keshet I, Siegfried Z, Mendelsohn M, Nemes A, Temper V, Razin A, Cedar H (1994) Sp1 element protect a CpG island from de novo methylation. Nature 371:435-438

Breathnach R and Chambon P (1981) Organization and expression of eukaryotic split genes coding for proteins. Annu Rev Biochem 50:349-383

Briggs MR, Kadonaga JT, Bell SP, Tjian R (1986) Purification and biochemical characterization of the promoter- specific transcription factor, Sp1. Science 234:47–52

Brockman JL and Schuler LA (2005) Prolactin signals via Stat5 and Oct-1 to the proximal cyclin D1 promoter. Mol Cell Endocrinol 239:45-53

Brugnera E, Xu L, Schaffner W, Arnosti DN (1992) POU-specific domain of Oct-2 factor confers 'octamer' motif DNA binding specificity on heterologous Antennapedia homeodomain. FEBS Lett 314:361-365

Bucher P (1990) Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. J Mol Biol 212:563-578

Buchman AR, Kimmerly WJ, Rine J, Kornberg RD (1988) Two DNA-binding factors recognize specific sequences at silencers, upstream activating sequences, autonomously replicating sequences, and telomeres in Saccharomyces cerevisiae. Mol Cell Biol 8:210-225

Buchman AR, Lue NF, Kornberg RD (1988) Connections between transcriptional activators, silencers, and telomeres as revealed by functional analysis of a yeast DNA-binding protein. Mol Cell Biol 8:5086-5099

Burke TW and Kadonaga JT (1996) Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes Dev 10:711-724

Burke TW and Kadonaga JT (1997) The downstream core promoter element, DPE, is conserved from Drosophila to human and is recognized by TAFII60 of Drosophila. Genes Dev 11:3020-3031

Burley SK and Roeder RG (1996) Biochemistry and structural biology of transcription factor IID (TFIID). Annu Rev Biochem 65:769-799

Burt AD (2004) Spectrum of liver pathology in immune mediated liver diseases. J Gastroenterol Hepatol. Suppl 9:S353-355

Butler JEF and Kadonaga JT (2002) The RNA polymerase II core promoter: a key component in the regulation of gene expression. Genes & Dev 16:2583-2592

Chambers A, Tsang JS, Stanway C, Kingsman AJ, Kingsman SM (1989) Transcriptional control of the Saccharomyces cerevisiae PGK gene by RAP1. Mol Cell Biol 9:5516-5524

Capieaux E, Vignais ML, Sentenac A, Goffeau A (1989) The yeast H+-ATPase gene is controlled by the promoter binding factor TUF. J Biol Chem 264:7437-7446

Carey M (1998) The enhanceosome and transcriptional synergy. Cell 92:5-8

Cawley S, Bekiranov S, Ng HH, Kapranov P, Sekinger EA, Kampa D, Piccolboni A, Sementchenko V, Cheng J, Williams AJ, Wheeler R, Wong B, Drenkow J, Yamanaka M, Patel S, Brubaker S, Tammana H, Helt G, Struhl K, Gingeras TR (2004) Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. Cell 116:499-509

Chandran UR, Warren BS, Baumann CT, Hager GL, DeFranco DB (1999) The glucocorticoid receptor is tethered to DNA-bound Oct-1 at the mouse gonadotropin-releasing hormone distal negative glucocorticoid response element. J Biol Chem 274:2372-2378

Chiba T, Goto S, Yokosuka O, Imazeki F, Tanaka M, Fukai K, Takahashi Y, Tsujimura H, Saisho H (2004) Fatal chronic active Epstein-Barr virus infection mimicking autoimmune hepatitis. Eur J Gastroenterol Hepatol 16:225-228

Comai L, Tanese N, Tjian R (1992) The TATA-binding protein and associated factors are integral components of the RNA polymerase I transcription factor, SL1. Cell 68:965-976

Corden J, Wasylyk B, Buchwalder A, Sassone-Corsi P, Kedinger C, Chambon P (1980) Promoter sequences of eukaryotic protein-coding genes. Science 209:1406-1414

Costa M, Rodriguez-Sanchez JL, Czaja AJ, Gelpi C (2000) Isolation and characterization of cDNA encoding the antigenic protein of the human tRNP(Ser)Sec complex recognized by autoantibodies from patients with type-1 autoimmune hepatitis. Clin Exp Immunol 121:364-374

Courey AJ, Holtzman DA, Jackson SP, Tjian R (1989) Synergistic activation by the glutamine-rich domains of human transcription factor Sp1. Cell 59:827–836

Cron RQ, Zhou B, Brunvand MW, Lewis DB (2001) Octamer proteins inhibit IL-4 gene transcription in normal human CD4 T cells. Genes Immun 2:464-468

Cross SH and Bird AP (1995) CpG islands and genes. Curr Opin Genet Dev 5:309-314

Czaja AJ, Souto EO, Bittencourt PL, Cancado EL, Porta G, Goldberg AC, Donaldson PT (2002) Clinical distinctions and pathogenic implications of type 1 autoimmune hepatitis in Brazil and the United States. J Hepatol 37:302-308

Devlin C, Tice-Baldwin K, Shore D, Arndt KT (1991) RAP1 is required for BAS1/BAS2- and GCN4-dependent transcription of the yeast HIS4 gene. Mol Cell Biol 11:3642-3651

Djilali-Saiah I, Renous R, Caillat-Zucman S, Debray D, Alvarez F (2004) Linkage disequilibrium between HLA class II region and autoimmune hepatitis in pediatric patients. J Hepatol 40:904-909

Dynan WS and Tjian R (1983) The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. Cell 35:79-87

Ellis J, Tan-Un KC, Harper A, Michalovich D, Yannoutsos N, Philipsen S, Grosveld F (1996) A dominant chromatin-opening activity in 5' hypersensitive site 3 of the human beta-globin locus control region. EMBO J 15:562-568

Emami KH, Navarre WW, Smale ST (1995) Core promoter specificities of the Sp1 and VP16 transcriptional activation domains. Mol Cell Biol 15:5906-5916

Fadel BM, Boutet SC, Quertermous T (1999), Octamer-dependent in vivo expression of the endothelial cell-specific TIE2 gene. J Biol Chem. 1999 274:20376-20383

Fantino E, Marguet D, Lauquin GJ (1992) Downstream activating sequence within the coding region of a yeast gene: specific binding in vitro of RAP1 protein. Mol Gen Genet 236:65-75

Fletcher C, Heintz N, Roeder RG (1987) Purification and characterization of OTF-1, a transcription factor regulating cell cycle expression of a human histone H2b gene. Cell 51:773-781

Franke S, Scholz G, Scheidereit C (1994) Identification of novel ubiquitous and cell type-specific factors that specifically recognize immunoglobulin heavy chain and kappa light chain promoters. J Biol Chem 269:20075-20082

Gardiner-Garden M and Frommer M (1987) CpG islands in vertebrate genomes. J Mol Biol 196:261-282

Gelpi C, Sontheimer EJ, Rodriguez-Sanchez JL (1992) Autoantibodies against a serine tRNA-protein complex implicated in cotranslational selenocysteine insertion. Proc Natl Acad Sci USA 89: 9739-43

Gershenzon NI and Ioshikhes IP (2005) Synergy of human Pol II core promoter elements revealed by statistical sequence analysis. Bioinformatics 21:1295-1300

Gidoni D, Dynan WS, Tjian R (1984) Multiple specific contacts between a mammalian transcription factor and its cognate promoters. Nature 312:409-413

Goodrich JA and Tjian R (1994) Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. Cell 77:145-156

Gough A, Chapman S, Wagstaff K, Emery P, Elias E (1996) Minocycline induced autoimmune hepatitis and systemic lupus erythematosus-like syndrome. BMJ 312:169-172

Grabe N (2002) AliBaba2: context specific identification of transcription factor binding sites. In Silico Biol. 2:S1-15

Graham IR and Chambers A (1994) Use of a selection technique to identify the diversity of binding sites for the yeast RAP1 transcription factor. Nucleic Acids Res 22:124-130

Graziadei IW, Obermoser GE, Sepp NT, Erhart KH, Vogel W (2003) Drug-induced lupus-like syndrome associated with severe autoimmune hepatitis. Lupus 12:409-412

Green MR (2000) TBP-associated factors (TAFIIs): multiple, selective transcriptional mediators in common complexes. Trends Biochem Sci 25:59-63

Grosveld F, van Assendelft GB, Greaves DR, Kollias G (1987) Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell 51:975-985

Gstaiger M, Georgiev O, van Leeuwen H, van der Vliet P, Schaffner W (1996) The B cell coactivator Bob1 shows DNA sequence-dependent complex formation with Oct-1/Oct-2 factors, leading to differential promoter activation. EMBO J 15:2781-2790

Hampsey M (1998) Molecular genetics of the RNA polymerase II general transcriptional machinery. Microbiol Mol Biol Rev 62:465-503

Harbison CT, Gordon DB, Lee TI, Rinaldi NJ, Macisaac KD, Danford TW, Hannett NM, Tagne JB, Reynolds DB, Yoo J, Jennings EG, Zeitlinger J, Pokholok DK, Kellis M, Rolfe PA, Takusagawa KT, Lander ES, Gifford DK, Fraenkel E, Young RA (2004) Transcriptional regulatory code of a eukaryotic genome. Nature 431:99-104

Henry YA, Chambers A, Tsang JS, Kingsman AJ, Kingsman SM (1990) Characterisation of the DNA binding domain of the yeast RAP1 protein. Nucleic Acids Res 18:2617-2623 Erratum in: Nucleic Acids Res 1990; 18:4317

Herkel J, Heidrich B, Nieraad N, Wies I, Rother M, Lohse AW (2002) Fine specificity of autoantibodies to soluble liver antigen and liver/pancreas. Hepatology 35:403-408

Herr W, Sturm RA, Clerc RG, Corcoran LM, Baltimore D, Sharp PA, Ingraham HA, Rosenfeld MG, Finney M, Ruvkun G, et al (1988) The POU domain: a large conserved region in the mammalian pit-1, oct-1, oct-2, and Caenorhabditis elegans unc-86 gene products. Genes Dev 2:1513-1516

Hirose T, Sowa Y, Takahashi S, Saito S, Yasuda C, Shindo N, Furuichi K, Sakai T (2003) p53-independent induction of Gadd45 by histone deacetylase inhibitor: coordinate regulation by transcription factors Oct-1 and NF-Y. Oncogene 22:7762-7773

Iademarco MF, McQuillan JJ, Rosen GD, Dean DC (1992) Characterization of the promoter for vascular cell adhesion molecule-1 (VCAM-1). J Biol Chem 267:16323-16329

Iademarco MF, McQuillan JJ, Dean DC (1993) Vascular cell adhesion molecule 1: contrasting transcriptional control mechanisms in muscle and endothelium. Proc Natl Acad Sci USA 90:3943-3947

Idrissi FZ, Fernandez-Larrea JB, Pina B (1998) Structural and functional heterogeneity of Rap1p complexes with telomeric and UASrpg-like DNA sequences. J Mol Biol 284:925-935

Idrissi FZ and Pina B (1999) Functional divergence between the half-sites of the DNA-binding sequence for the yeast transcriptional regulator Rap1p. Biochem J 341:477-482

Imai S, Nishibayashi S, Takao K, Tomifuji M, Fujino T, Hasegawa M, Takano T (1997) Dissociation of Oct-1 from the nuclear peripheral structure induces the cellular aging-associated collagenase gene expression. Mol Biol Cell 8:2407-2419

Jackson SP and Tjian R (1988) O-glycosylation of eukaryotic transcription factors: implications for mechanisms of transcriptional regulation. Cell 55:125–133

Jackson S, MacDonald JJ, Lees-Miller S, Tjian R (1990) GC box binding induces phosphorylation of Sp1 by a DNA-dependent protein kinase. Cell 63:155–165

Janson L and Pettersson U (1990) Cooperative interactions between transcription factors Sp1 and OTF-1. Proc Natl Acad Sci USA 87:4732-4736

Javahery R, Khachi A, Lo K, Zenzie-Gregory B, Smale ST (1994) DNA sequence requirements for transcriptional initiator activity in mammalian cells. Mol Cell Biol 14:116-127

Jensen DE, Black AR, Swick AG, Azizkhan JC (1997) Distinct roles for Sp1 and E2F sites in the growth/cell cycle regulation of the DHFR promoter. J Cell Biochem 67:24-31

Jin S, Fan F, Fan W, Zhao H, Tong T, Blanck P, Alomo I, Rajasekaran B, Zhan Q (2001) Transcription factors Oct-1 and NF-YA regulate the p53-independent induction of the GADD45 following DNA damage. Oncogene 20:2683-2690

Jongstra J, Reudelhuber TL, Oudet P, Benoist C, Chae CB, Jeltsch JM, Mathis DJ, Chambon P (1984) Induction of altered chromatin structures by simian virus 40 enhancer and promoter elements. Nature 307:708-714

Kaczynski J, Cook T, Urrutia R (2003) Sp1- and Kruppel-like transcription factors. Genome Biol 4:206

Kadonaga JT, Carner KR, Masiarz FR, Tjian R (1987) Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. Cell 51:1079-1090

Kadonaga JT (1990) Assembly and disassembly of the Drosophila RNA polymerase II complex during transcription. J Biol Chem 265:2624-2631

Kadonaga JT (2002) The DPE, a core promoter element for transcription by RNA polymerase II. Exp Mol Med 34:259-264

Kakizawa T, Miyamoto T, Ichikawa K, Kaneko A, Suzuki S, Hara M, Nagasawa T, Takeda T, Mori J, Kumagai M, Hashizume K (1999) Functional interaction between Oct-1 and retinoid X receptor. J Biol Chem 274:19103-19108

Kakizawa T, Miyamoto T, Ichikawa K, Takeda T, Suzuki S, Mori J, Kumagai M, Yamashita K, Hashizume K (2001) Silencing mediator for retinoid and thyroid hormone receptors interacts with octamer transcription factor-1 and acts as a transcriptional repressor. J Biol Chem 276:9720-9725

Karlseder J, Rotheneder H, Wintersberger E (1996) Interaction of Sp1 with the growth- and cell cycle-regulated transcription factor E2F. Mol Cell Biol 16:1659-1667

Kavurma MM, Santiago FS, Bonfoco E, Khachigian LM (2001) Sp1 phosphorylation regulates apoptosis via extracellular FasL-Fas engagement. J Biol Chem 276:4964-4971

Kavurma MM and Khachigian LM (2003) Sp1 inhibits proliferation and induces apoptosis in vascular smooth muscle cells by repressing p21WAF1/Cip1 transcription and cyclin D1-Cdk4-p21WAF1/Cip1 complex formation. J Biol Chem 278:32537-32543

Kernebeck T, Lohse AW, Grotzinger J (2001) A bioinformatical approach suggests the function of the autoimmune hepatitis target antigen soluble liver antigen/liver pancreas. Hepatology 34:230-233

Kim MH and Peterson DO (1995) Stimulation of basal transcription from the mouse mammary tumor virus promoter by Oct proteins. J Virol 69:4717-4726

Konig P, Giraldo R, Chapman L, Rhodes D (1996) The crystal structure of the DNA-binding domain of yeast RAP1 in complex with telomeric DNA. Cell 85:125-136

Konig P and Rhodes D (1997) Recognition of telomeric DNA. Trends Biochem Sci 22:43-47

Krawitt EL (1996) Autoimmune hepatitis. N Engl J Med 334:897-903

Krawitt EL (2006) Medical progress: autoimmune hepatitis. N Engl J Med 354:54-66

Kutach AK and Kadonaga JT (2000) The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoter. Mol Cell Biol 20:4754-4764

Kutoh E, Stromstedt PE, Poellinger L (1992) Functional interference between the ubiquitous and constitutive octamer transcription factor 1 (OTF-1) and the glucocorticoid receptor by direct protein-protein interaction involving the homeo subdomain of OTF-1. Mol Cell Biol 12:4960-4969

Labrador M and Corces VG (2002) Setting the boundaries of chromatin domains and nuclear organization. Cell 111:151-154

Lagrange T, Kapanidis AN, Tang H, Reinberg D, Ebright RH (1998) New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. Genes & Dev 12:34-44

Larsen F, Gundersen G, Lopez R, Prydz H (1992) CpG islands as gene markers in the human genome. Genomics 13:1095-1107

Laskus T and Slusarczyk J (1989) Autoimmune chronic active hepatitis developing after acute type B hepatitis. Dig Dis Sci 34:1294–1297

Lemon B and Tjian R (2000) Orchestrated response: a symphony of transcription factors for gene control. Genes & Dev 14:2551-2569

Lewis BA, Kim TK, Orkin SH (2000) A downstream element in the human beta-globin promoter: evidence of extended sequence-specific transcription factor IID contacts. Proc Natl Acad Sci USA 97:7172-7177

Li B, Oestreich S, de Lange T (2000) Identification of human Rap1: implications for telomere evolution. Cell 101:471-483

Li L, He S, Sun JM, Davie JR (2004) Gene regulation by Sp1 and Sp3. Biochem Cell Biol 82:460-471

Li R, Knight JD, Jackson SP, Tjian R, Botchan MR (1991) Direct interaction between Sp1 and the BPV enhancer E2 protein mediates synergistic activation of transcription. Cell 65:493–505

Lim CY, Santoso B, Boulay T, Dong E, Ohler U, Kadonaga JT (2004) The MTE, a new core promoter element for transcription by RNA polymerase II. Genes Dev 18:1606-1617

Li-Weber M, Laur O, Hekele A, Coy J, Walczak H, Krammer PH (1998) A regulatory element in the CD95 (APO-1/Fas) ligand promoter is essential for responsiveness to TCR-mediated activation. Eur J Immunol 28:2373-2383

Lo K and Smale ST (1996) Generality of a functional initiator consensus sequence. Gene 182:13-22

Lodish H (2000) Molecular cell biology, 4_{th} edn. WH Freeman & Company, New York

Luo Y, Fujii H, Gerster T, Roeder RG (1992) A novel B cell-derived coactivator potentiates the activation of immunoglobulin promoters by octamer-binding transcription factors. Cell 71:231-241

Luo Y and Roeder RG (1995) Cloning, functional characterization, and mechanism of action of the B-cell-specific transcriptional coactivator OCA-B. Mol Cell Biol 15:4115-4124

Macleod D, Charlton J, Mullins J, Bird AP (1994) Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. Genes Dev 8:2282-2292

Marin M, Karis A, Visser P, Grosveld F, Philipsen S (1997) Transcription factor Sp1 is essential for early embryonic development but dispensable for cell growth and differentiation. Cell 89:619-628

McClure RF, Heppelmann CJ, Paya CV (1999) Constitutive Fas ligand gene transcription in Sertoli cells is regulated by Sp1. J Biol Chem 274:7756-7762

Magne S, Caron S, Charon M, Rouyez MC, Dusanter-Fourt I (2003) STAT5 and Oct-1 form a stable complex that modulates cyclin D1 expression. Mol Cell Biol 23:8934-8945

Martinez E, Zhou Q, L'Etoile ND, Oelgeschlager T, Berk AJ, Roeder RG (1995) Core promoter-specific function of a mutant transcription factor TFIID defective in TATA-box binding. Proc Natl Acad Sci USA 92:11864-11868

Mesplede T, Island ML, Christeff N, Petek F, Doly J, Navarro S (2005) The POU transcription factor Oct-1 represses virus-induced interferon A gene expression. Mol Cell Biol 25:8717-8731

Medina J, García-buey L, Moreno-Otero R (2003) Review article: immunopathogenetic and therapeutic aspects of autoimmune hepatitis. Aliment Pharmacol Ther 17:1-16

Morse RH (2000) RAP, RAP, open up! New wrinkles for RAP1 in yeast. Trends Genet 16:51-53

Murphy S, Pierani A, Scheidereit C, Melli M, Roeder RG (1989) Purified octamer binding transcription factors stimulate RNA polymerase III--mediated transcription of the 7SK RNA gene. Cell 59:1071-1080

Nakshatri H, Nakshatri P, Currie RA (1995) Interaction of Oct-1 with TFIIB. Implications for a novel response elicited through the proximal octamer site of the lipoprotein lipase promoter. J Biol Chem 270:19613-19623

Nieuwint RT, Mager WH, Maurer KC, Planta RJ (1989) Mutational analysis of the upstream activation site of yeast ribosomal protein genes. Curr Genet 15:247-251

Nikolov DB and Burley SK (1997) RNA polymerase II transcription initiation: a structural view. Proc Natl Acad Sci USA 94:15-22

Ohler U, Liao GC, Niemann H, Rubin GM (2002) Computational analysis of core promoters in the Drosophila genome. Genome Biol 3:0087.1-0087.12

Oren T, Torregroza I, Evans T (2005) An Oct-1 binding site mediates activation of the gata2 promoter by BMP signaling. Nucleic Acids Res 33:4357-4367

Osborne A, Zhang H, Yang WM, Seto E, Blanck G (2001) Histone deacetylase activity represses gamma interferon-inducible HLA-DR gene expression following the establishment of a DNase I-hypersensitive chromatin conformation. Mol Cell Biol 21:6495-6506

Osborne AR, Zhang H, Fejer G, Palubin KM, Niesen MI, Blanck G (2004) Oct-1 maintains an intermediate, stable state of HLA-DRA promoter repression in Rb-defective cells: an Oct-1-containing repressosome that prevents NF-Y binding to the HLA-DRA promoter. J Biol Chem 279:28911-28919

O'Shea-Greenfield A and Smale ST (1992) Roles of TATA and initiator elements in determining the start site location and direction of RNA polymerase II transcription. J Biol Chem 267:1391-1402 Erratum in: J Biol Chem 1992; 267:6450.

Palladino F, Laroche T, Gilson E, Pillus L, Gasser SM (1993) The positioning of yeast telomeres depends on SIR3, SIR4, and the integrity of the nuclear membrane. Cold Spring Harb Symp Quant Biol 58:733-746

Pando M, Larriba J, Fernandez GC, Fainboim H, Ciocca M, Ramonet M, Badia I, Daruich J, Findor J, Tanno H, Canero-Velasco C, Fainboim L (1999) Pediatric and adult forms of type I autoimmune hepatitis in Argentina: evidence for differential genetic predisposition. Hepatology 30:1374-1380

Pankratova EV, Deyev IE, Zhenilo SV, Polanovsky OL (2001) Tissue-specific isoforms of the ubiquitous transcription factor Oct-1. Mol Genet Genomics 266:239-245

Park SZ, Nagorney DM, Czaja AJ (2000) Hepatocellular carcinoma in autoimmune hepatitis. Dig Dis Sci 45:1944-1948

Ponger L and Mouchiroud D (2001) CpGProD: identifying CpG islands associated with transcription start sites in large genomic mammalian sequences. Bioinformatics 18:631-633

Prabhu A, O'Brien DP, Weisner GL, Fulton R, Van Ness B (1996) Octamer independent activation of transcription from the kappa immunoglobulin germline promoter. Nucleic Acids Res 24:4805-4811

Prefontaine GG, Lemieux ME, Giffin W, Schild-Poulter C, Pope L, LaCasse E, Walker P, Hache RJ (1998) Recruitment of octamer transcription factors to DNA by glucocorticoid receptor. Mol Cell Biol 18:3416-3430

Pugh BF and Tjian R (1990) Mechanism of transcriptional activation by Sp1: evidence for oactivators. Cell 61:1187-1197

Roberts SB, Segil N, Heintz N (1991) Differential phosphorylation of the transcription factor Oct1 during the cell cycle. Science 253:1022-1026

Robertson DA, Zhang SL, Guy EC, Wright R (1987) Persistent measles virus genome in autoimmune chronic active hepatitis. Lancet 2:9-11

Rosenfeld MG (1991) POU-domain transcription factors: pou-er-ful developmental regulators. Genes Dev 5:897-907

Rundlof AK, Carlsten M, Arner ES (2001) The core promoter of human thioredoxin reductase 1: cloning, transcriptional activity, and Oct-1, Sp1, and Sp3 binding reveal a housekeeping-type promoter for the AU-rich element-regulated gene. J Biol Chem 276:30542-30551

Ruvkun G and Finney M (1991) Regulation of transcription and cell identity by POU domain proteins. Cell 64:475-478

Ryan AK and Rosenfeld MG (1997) POU domain family values: flexibility, partnerships, and developmental codes. Genes Dev 11:1207-1225

Scarsdale JN, Kazanina G, Radaev S, Schirch V, Wright HT (1999) Crystal structure of rabbit cytosolic serine hydroxymethyltransferase at 2.8 A resolution: mechanistic implications. Biochemistry 38:8347-8358

Schild-Poulter C, Shih A, Yarymowich NC, Hache RJ (2003) Down-regulation of histone H2B by DNA-dependent protein kinase in response to DNA damage through modulation of octamer transcription factor 1. Cancer Res 63:7197-7205

Scholer HR (1991) Octamania: the POU factors in murine development. Trends Genet 7:323-329

Schwachtgen JL, Remacle JE, Janel N, Brys R, Huylebroeck D, Meyer D, Kerbiriou-Nabias D (1998) Oct-1 is involved in the transcriptional repression of the von willebrand factor gene promoter. Blood 92:1247-1258

Shore D, Stillman DJ, Brand AH, Nasmyth KA (1987) Identification of silencer binding proteins from yeast: possible roles in SIR control and DNA replication. EMBO J 6:461-467

Shore D and Nasmyth K (1987) Purification and cloning of a DNA binding protein from yeast that binds to both silencer and activator elements. Cell 51:721-732

Shore D (1994) RAP1: a protean regulator in yeast. Trends Genet 10:408-412
Skoog SM, Rivard RE, Batts KP, Smith CI (2002) Autoimmune hepatitis preceded by acute hepatitis A infection. Am J Gastroenterol 97:1568-1569

Smale ST and Baltimore D (1989) The “initiator” as a transcription control element. Cell 57:103-113

Smale ST, Schmidt MC, Berk AJ, Baltimore D (1990) Transcriptional activity by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID. Proc Natl Acad Sci USA 87:4509-4513

Smale ST, Jain A, Kaufmann J, Emami KH, Lo K, Garraway IP (1998) The initiator element: a paradigm for core promoter heterogeneity within metazoan protein-coding genes. Cold Spring Harb Symp Quant Biol 63:21-31

Smale ST (2001) Core promoters: active contributors to combinatorial gene regulation. Genes & Dev 15:2503-2508

Smale ST and Kadonaga JT (2003) The RNA polymerase II core promoter. Annu Rev Biochem 72:449-479

Sterling MJ, Kane M, Grace ND (1996) Pemoline-induced autoimmune hepatitis. Am J Gastroenterol 91:2233-2234

Strom AC, Forsberg M, Lillhager P, Westin G (1996) The transcription factors Sp1 and Oct-1 interact physically to regulate human U2 snRNA gene expression. Nucleic Acids Res 24:1981-1986

Strubin M, Newell JW, Matthias P (1995) OBF-1, a novel B cell-specific coactivator that stimulates immunoglobulin promoter activity through association with octamer-binding proteins. Cell 80:497-506

Sturm R, Baumruker T, Franza BR Jr, Herr W (1987) A 100-kD HeLa cell octamer binding protein (OBP100) interacts differently with two separate octamer-related sequences within the SV40 enhancer. Genes Dev 1:1147-1160

Sturm RA and Herr W (1988) The POU domain is a bipartite DNA-binding structure. Nature 336:601-604

Sturm RA, Das G, Herr W (1988) The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. Genes Dev 2:1582-1599

Subramaniam N, Cairns W, Okret S (1998) Glucocorticoids repress transcription from a negative glucocorticoid response element recognized by two homeodomain-containing proteins, Pbx and Oct-1. J Biol Chem 273:23567-23574

Suske G (1999) The Sp-family of transcription factors. Gene 238:291-300

Suzuki Y, Tsunoda T, Sese J, Taira H, Mizushima-Sugano J, Hata H, Ota T, Isogai T, Tanaka T, Nakamura Y, Suyama A, Sakaki Y, Morishita S, Okubo K, Sugano S (2001) Identification and characterization of the potential promoter regions of 1031 kinds of human genes. Genome Research 11:677-684
Swick AG, Blake MC, Kahn JW, Azizkhan JC (1989) Functional analysis of GC element binding and transcription in the hamster dihydrofolate reductase gene promoter. Nucleic Acids Res 17:9291-9304

Tanese N, Pugh BF, Tjian R (1991) Coactivators for a proline-rich activator purified from the multisubunit human TFIID complex. Genes Dev 5:2212-2224

Tanikawa J, Yasukawa T, Enari M, Ogata K, Nishimura Y, Ishii S, Sarai A (1993) Recognition of specific DNA sequences by the c-myb protooncogene product: role of three repeat units in the DNA-binding domain. Proc Natl Acad Sci USA 90:9320-9324

Tsai FT and Sigler PB (2000) Structural basis of preinitiation complex assembly on human pol II promoters. EMBO J 19:25-36

Toda G, Zeniya M, Watanabe F, Imawari M, Kiyosawa K, Nishioka M, Tsuji T, Omata M (1997) Present status of autoimmune hepatitis in Japan – correlating the characteristics with international criteria in an area with a high rate of HCV infection. J Hepatol 26:1207-1212

Tornow J and Santangelo GM (1990) Efficient expression of the Saccharomyces cerevisiae glycolytic gene ADH1 is dependent upon a cis-acting regulatory element (UASRPG) found initially in genes encoding ribosomal proteins. Gene 90:79-85

Valenzuela P, Bell GI, Weinberg F, Rutter WJ (1976) Yeast DNA dependent RNA polymerase I, II and III. The existence of subunits common to the enzymes. Biochem Biophys Res Commun 71:1319-1325

Vento S, Cainelli F, Renzini C, Concia E (1997) Autoimmune hepatitis type 2 induced by HCV and persisting after viral clearance. Lancet 350:1298-1299

Verrijzer CP, Alkema MJ, van Weperen WW, Van Leeuwen HC, Strating MJ, van der Vliet PC (1992) The DNA binding specificity of the bipartite POU domain and its subdomains. EMBO J 11:4993-5003

Verrijzer CP and Van der Vliet PC (1993) POU domain transcription factors. Biochim Biophys Acta 1173:1-21

Verrijzer CP, Chen J-L, Yokomori K, Tjian R (1995) Binding of TAFs to core elements directs promoter selectivity by RNA polrmersae II. Cell 81:1115-1125

Voss JW, Wilson L, Rosenfeld MG (1991) POU-domain proteins Pit-1 and Oct-1 interact to form a heteromeric complex and can cooperate to induce expression of the prolactin promoter. Genes Dev 5:1309-1320

Wahlin J and Cohn M (2000) Saccharomyces cerevisiae RAP1 binds to telomeric sequences with spatial flexibility. Nucleic Acids Res 28:2292-2301

Wang CX, Teufel A, Cheruti U, Grotzinger J, Galle PR, Lohse AW, Herkel J (2006) Characterisation of human gene encoding SLA/LP autoantigen and its conserved homologs in mouse, fish, fly and worm. World J Gastroenterol 12:902-907

Wang JC and Van Dyke MW (1993) Initiator sequences direct downstream promoter binding by human transcription factor IID. Biochim Biophys Acta 1216:73-80

Warrington JA, Nair A, Mahadevappa M, Tsyganskaya M (2000) Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes. Physiol Genomics 2:143-147

West AG, Gaszner M, Felsenfeld G (2002) Insulators: many functions, many mechanisms. Genes Dev 16:271-288

Wies I, Brunner S, Henninger J, Herkel J, Kanzler S, Meyer zum Buschenfelde KH, Lohse AW (2000) Identification of target antigen for SLA/LP autoantibodies in autoimmune hepatitis. Lancet 355:1510-1515

Witt O, Albig W, Doenecke D (1997) Transcriptional regulation of the human replacement histone gene H3.3B. FEBS Lett 408:255-260

Woudt LP, Mager WH, Nieuwint RT, Wassenaar GM, van der Kuyl AC, Murre JJ, Hoekman MF, Brockhoff PG, Planta RJ (1987) Analysis of upstream activation sites of yeast ribosomal protein genes. Nucleic Acids Res 15:6037-6048

Wu GD, Lai EJ, Huang N, Wen X (1997) Oct-1 and CCAAT/enhancer-binding protein (C/EBP) bind to overlapping elements within the interleukin-8 promoter. The role of Oct-1 as a transcriptional repressor. J Biol Chem 272:2396-2403

Xu XM, Mix H, Carlson BA, Grabowski PJ, Gladyshev VN, Berry MJ, Hatfield DL (2005) Carlson, et al. Evidence for Direct Roles of Two Additional Factors, SECp43 and Soluble Liver Antigen, in the Selenoprotein Synthesis Machinery. The Journal of Biological Chemistry 280:41568-41575

Yoshizawa K, Ota M, Katsuyama Y, Ichijo T, Matsumoto A, Tanaka E, Kiyosawa K (2005). Genetic analysis of HLA region of Japanese patients with type 1 autoimmune hepatitis. J Hepatol 42:578-584

Yu B, Datta PK, Bagchi S (2003) Stability of the Sp3-DNA complex is promoter-specific: Sp3 efficiently competes with Sp1 for binding to promoters containing multiple Sp-sites. Nucleic Acid Res 31:5368-5376

Zenzie-Gregory B, Khachi A, Garraway IP, Smale ST (1993) Mechanism of initiator-mediated transcription: evidence for a functional interaction between the TATA-binding protein and DNA in the absence of a specific recognition sequence. Mol Cell Biol 13:3841-3849

Zenzie-Gregory B, O'Shea-Greenfield A, Smale ST (1992) Similar mechanisms for transcription initiation mediated through a TATA box or an initiator element. J Biol Chem 267:2823-2830

Zhao FQ, Zheng Y, Dong B, Oka T (2004) Cloning, genomic organization, expression, and effect on beta-casein promoter activity of a novel isoform of the mouse Oct-1 transcription factor. Gene 326:175-187

Zhang H, Shepherd AT, Eason DD, Wei S, Diaz JI, Djeu JY, Wu GD, Blanck G (1999) Retinoblastoma protein expression leads to reduced Oct-1 DNA binding activity and enhances interleukin-8 expression. Cell Growth Differ 10:457-465

Zhang L, Li WH (2004) Mammalian housekeeping genes evolve more slowly than tissue-specific genes. Mol Biol Evol. 21(2):236-239

Zhao H, Jin S, Fan F, Fan W, Tong T, Zhan Q (2000) Activation of the transcription factor Oct-1 in response to DNA damage. Cancer Res 60:6276-6280

Zwilling S, Annweiler A, Wirth T (1994) The POU domains of the Oct1 and Oct2 transcription factors mediate specific interaction with TBP. Nucleic Acids Res 22:1655-1662

Zwilling S, Konig H, Wirth T (1995) High mobility group protein 2 functionally interacts with the POU domains of octamer transcription factors. EMBO J 14:1198-1208

7. Abbreviations

AIH	Autoimmune hepatitis
ANA	Antinuclear antibody
ATP	Adenosine 5'-triphosphate
BAC	Bacterial artificial chromosome
bp	Base pair
BRE	Recognition element
bZIP	basic zipper
bHLH	basic helix-loop-helix
CIP	Calf Intestinal Alkaline Phosphatase
CpG	Dinucleotide of C and G
Cys	Cysteine
DCE	Downstream core element
dCTP	Deoxycytidine triphosphate
DMEM	Dulbecco's Modified Eagle Medium
DPE	Downstream promoter element
EMSA	Electrophoretic mobility shift assay
FCS	Fetal Calf Serum
Fw	Forward
HEK293	Human embryonic kidney epithelial cell line
HeLa	Human epithelial carcinoma cell line
Hep1-6	Mouse hepatocellular cell line
His	Histidine
HLA	Human leukocyte antigen
HMG2	High mobility group protein 2
HMRE	Mating type silencer
Inr	Initiator
kb	Kilobase pair
kDa	Kilodalton
LC-1	Liver cytosol-1 antibody
LKM	Anti liver-kidney microsome-1 antibody
MAT	Mating type information locus
MTE	Motif ten element

Myb	Myeloblastosis oncogene
nt	nucleotide
OBF-1	Oct-binding factor-1
OCA-B	Oct-1-associated coactivator
Oct-1	Octamer transcription factor-1
pANCA	Perinuclear anti-neutrophilic cytoplasmic antibody
PBC	Primary biliary cirrhosis
PCR	Polymerase chain reaction
PIC	Pre-initiation complex
PLP	Pyridoxal phosphate
POU	Abbreviation of Pit-1, Oct-1/Oct-2 and Unc-86
PSC	Primary sclerosing cholangitis
PVDF	Polyvinylidene fluoride
Py	Pyrimidine
RAP1	Repressor activator protein 1
RAW264.7	Murine macrophage cell line
Re	Reverse
RLB	Reporter Lysis Buffer
RT	Reverse Transcription
SAP	Shrimp Alkaline Phosphatase
SECp43	A selenocysteine tRNA associated RNA-binding protein clone
Sir	Silent information regulator
SLA/LP	soluble liver antigen/liver pancreas antigen
SMA	Smooth-muscle antibody
Sp1	Stimulating protein 1 or Specificity protein 1
TAF	Transcription associated factor
TAZ1	Transcriptional adaptor zinc-binding domain
TBP	TATA box-binding protein
TSS	Transcription start site
TF	Transcription factor
Tm	Melting temperature
TRF1	Telomeric repeat binding factor 1
tRNP	tRNA-associated protein
(UAS)rpg	Upstream activator sequence

8. Acknowledgments

This study was conducted with the help of many people. First, I want give my special thanks to Prof. Dr. Ansgar W. Lohse, my supervisor, who offered me this opportunity to do this research in his group. I knew him six years ago due to his famous work and publication about SLA/LP. I had the idea to work and study in his group from that on. It came true in 2003 with the help of our friend, Shaojing Zhang who was a very nice man and will live in my memory forever. Prof. Lohse is an erudite and fascinating group leader, and he always arranges everything for me in advance that make me possible to concentrate on my work. I enjoy the time working in his group.

I give my thanks to Dr. Johannes Herkel and Dr. Andreas Teufel. They helped me with the project in detail, from the design of this research, suggestion of solving problems to the final editing of this dissertation. I learned quite a lot from them not only the technique but also the way of thinking. Within our attractive group, from Mainz to Hamburg, people always help each other without any hesitation. I am sure I will be reluctant to leave this group.

I want say thanks to my parents from the deep of my heart. I could not stay in Germany without their help. They took care of my son for four years. I can image the great effort they have made to insure my son healthy and study well.

I know that my husband, a constant supporter of me, suffered a lot during these years. I also feel sorry to my son that I did not stay with him during the very important teenager period. However, at the same time, I feel proud that I have so nice family that everybody did their best to support me. The only thing I can do to compensate them is to work hard to get good result as my mother always asked me how about your work even though she does not know it well.

Thanks to Shandong Provincial Hospital for giving me the permission to do research in Germany.

Finally, I want give my thanks to all the people who helped me. I can't write down all of their names because there are so many, Uta, Tanja, Christian, Paco, Katrin, Mark, Marko……

9. Curriculum Vitae

Name: Chunxia Wang

E-Mail: chunxiawang@gmx.de

Education:

2003 - 2006 Johannes Gutenberg University Mainz and University Medical Center of Hamburg-Eppendorf, research work about SLA/LP. Germany

2000 - 2002 Shandong University, Master of Medicine. China

1982 - 1987 Taishan Medical College, Bachelor of Medicine. China

Experiences:

01.2003 - 12.2006 Molecular biology research work in I Department of Medicine, Johannes Gutenberg University, Mainz and University Medical Center of Hamburg-Eppendorf , Mainz and Hamburg, Germany

03.1996 - 12.2002 Doctor in Charge and Associate Professor, Central Laboratory of Shandong Provincial Hospital, Jinan, China

10.1999 - 01.2000 Studied Immunofluorescence techniques in Euroimmune Company, Lübeck, Germany

07.1987 - 02.1996 Biochemistry teacher, Shandong Health School, Jinan, China

Publications:

1. Chunxia Wang, Andreas Teufel, Ansgar W. Lohse, Johannes Herkel. Identification and characterization of the promoter region of the SLA/LP gene. Submit.

2. Chunxia Wang, Andreas Teufel, Uta Cheruti, et al. Characterisation of human gene encoding SLA/LP autoantigen and its conserved homologues in mouse, fish, fly and worm. *World Journal of Gastroenterology*, 2006; 12(6):902

3. Chunxia Wang，Xiong Zou，Lanfang Wang，et al. Expression of MHC antigens in tissues and peripheral blood lymphocytes of patients with gastric carcinoma. *Journal of Tumor Marker Oncology*, 2003; 18 (1):72-79

4. Huaichen Li, Chunxia Wang, Chen Huang, et al. Effect and mechanism of arsenic trioxide on chemosensitivity of human lung adenocarcinoma cells. *Chinese Journal of Tuberculosis and Respiratory diseases*, 2003; 26 (11):689

5. Chunxia Wang, Hongyun Wang, Huaichen Li, et al. Expression of TGF-β on lung carcinoma. *ACTA Academiae Medicinae Shandong*, 2001;39 (2): 119

10. Statement

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe.

Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Unterschrift: ....