Characterization of Regulatory Mechanisms for Alternative Splicing in Alpha Thyroid Hormone Receptor mRNA

Chao Zhang
Marquette University

CHARACTERIZATION OF REGULATORY MECHANISMS FOR ALTERNATIVE SPLICING IN ALPHA THYROID HORMONE RECEPTOR mRNA

by

Chao Zhang, B.S

A Thesis submitted to the Faculty of the Graduate School,
Marquette University,
in Partial Fulfillment of the Requirements for
the Degree of Master of Science

Milwaukee, Wisconsin

August 2010

ABSTRACT
CHARACTERIZATION OF REGULATORY MECHANISMS FOR ALTERNATIVE SPLICING IN ALPHA THYROID HORMONE RECEPTOR mRNA

Chao Zhang, B.S

Marquette University, 2010

In eukaryotes, alternative splicing is an essential post transcriptional modification process for functional gene expression and a major contributor to protein diversity. The regulation of alternative splicing generally involves the engagements of RNA sequences cis-acting elements) and corresponding protein factors (trans-acting factors). The cis-acting RNA motifs can be categorized depending on positional and functional differences. Trans-acting protein factors will then bind to RNA sequences and affect the corresponding splicing activity. Recently, factors associated with 3' polyadenylation have also been identified to affect alternative splicing.

In mammals, the α-thyroid hormone receptor gene (TRα) produces transcripts for two functionally antagonistic isoforms, TRα1 and TRα2 by alternative splicing of pre-mRNA. TRα1 will activate the thyroid hormone responsive genes after the binding of thyroid hormone while TRα2, a non-hormone binding variant, plays a functionally antagonistic role. A third minor isoform, TRα3 has also been described which is similar to TRα2 but lacks part of the terminal sequence. Regulation of TRα alternative splicing requires the interaction of cis-trans elements and alternative polyadenylation. The goal of my project is to study the regulatory mechanism of TRα alternative splicing.

In our system, ESX10 (exonic splicing enhancer on exon10) was previously identified as a 200 nt splicing enhancer element located on the last exon of TRα gene. My study further characterized the enhancing capability of exonic splicing enhancers (ESE) motifs within ESX10. Three heptanucleotide ESE motifs in the 3' half of ESX10 have been identified. *In vitro* experiments indicate that substitution of eight nucleotides within three heptamers decreases the enhancing capability of ESX10. After substitution of ESX10 and its subfragments with Rev-erbβ sequence, cryptic splicing was detected and quantitated. Previous experiments also indicate that replacing original TRα1 and TRα2 poly (A) site with a strong SVL poly (A) signal would increase the corresponding mRNA expression. My results show that introduction of downstream poly (A) signal increases splicing fidelity. When upstream 5'ss of TRα2 is disabled, the strong poly (A) signal enhances the usage of weak splicing cites and promote cryptic splicing.

ACKNOWLEDGEMENTS

Chao Zhang, B.S

I would like to thank Dr.Munroe for his serious scientific attitude, noble personality and professional excellence. Without his encouragement and guidance, it will be impossible for me to finish this project and thesis. I would also like to thank Dr.James Anderson and Dr. Mark McNally for their help and being in my committee member for all three years. The suggestions from Dr.Rosemary Stuart and Dr.Michael Schlappi were also very insightful. These three years studying here will be a precious memory for my life that I would never forget.

Thanks to all the previous members and undergraduates worked in our lab. I appreciate the work of Dr. Michele Hasting, Ms.Valerie Salato and Ms Nathaniel W. Rediske for their diligent efforts where most of my work based on. Brandon Randflesch, Scott Brown, Corey Nemec and Tessa Hoff spent a long time in the lab and helped me in almost every aspect. I hope all of them have a bright and promising future in the academic society like my advisor.

I also appreciate the help and words from all the faculties, staff and other graduate students. They were friendly and kind all the time. Lastly, I would like to thank my parents for their support and motivation when I made my decision to continue my education and left for such long period.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter I. Introduction

A. Background

Alternative splicing in α-thyroid hormone (T3) receptor gene

In higher eukaryotes, the expression of genes requires accurate and complete transcription to produce functional pre-mRNAs. This primary transcripts is typically modified by alternative splicing, a process in which spliced exons are connected after intron removal [1, 2]. Most variations produced by alternative splicing are from splicing at different splicing sites within mRNAs from a single gene, which normally leads to expression of different protein isoforms. Since its discovery about 30 years ago, alternative splicing has been addressed as an essential part of post transcriptional modification and a major contributor to protein diversity from the genome [3, 4]. It has been estimated that 95–100% of human genes with more than one exon undergo alternative processing, which may account for the existence of far fewer protein coding genes than originally predicted [5, 6, 7, 8]. Mistakes in alternative splicing are associated with autoimmune disease, neurodegenerative disorders and various cancers [7, 8].

In mammals, the α-thyroid hormone (T3) receptor gene (TRα, also known as NR1A1) produces transcripts for two functional antagonistic protein factors, TRα1 and TRα2 through alternative splicing of TRα pre-mRNA [9,10,11,12]. The mature mRNAs for these two isoforms are identical except their 3' exons: TRα1 contains the complete sequence for the first 9 exons while the production of TRα2 requires specific splicing within exon 9 (at TRα2 specific 5' splice site) to exon 10 transcriptions (Figure 1). TRα1 mRNA encodes the authentic thyroid hormone receptor that activates the downstream

thyroid hormone (T3) responsive genes after binding T3. In contrast, translation of TRα2 produce a non-hormone binding variant that competes for the same DNA binding domain (T3 Response Elements) and will not involve in the following activation (Figure 1) [11,12]. TRα3, a further spliced isoform of TRα, has also been identified in rat from TRα alternative splicing. In TRα3, the proximal 39 amino acids sequence from exon10 of TRα2 has been deleted. Functionally, TRα3 is similar to TRα2 in that TRα3 receptor doesn't bind to thyroid hormone and won't activate downstream genes [13, 14].

In addition to the TRα1 and TRα2 mRNAs, mRNA for another nuclear receptor protein, Rev-erbα is also transcribed from overlapping region but in the opposite direction [15, 16]. Rev-erbα has been confirmed to play an important role in circadian regulation by controlling the transcription of BMAL1, one critical component of the molecular apparatus that regulates circadian rhythms [17, 18]. Rev-erbα was originally identified as an 'orphan receptor', which is a protein receptor that doesn't have a known ligand. However, recent studies demonstrated that heme reversibly binds to Rev-erbα in its canonical ligand binding domain and regulate its interaction with other regulatory proteins [17, 18]. Since TRα and Rev-erbα share a 263 nt exon-exon overlap at their 3' ends and their respective transcripts may interact with each other, the relationship between the sequence of these two physiological important genes has been studied by our laboratory.

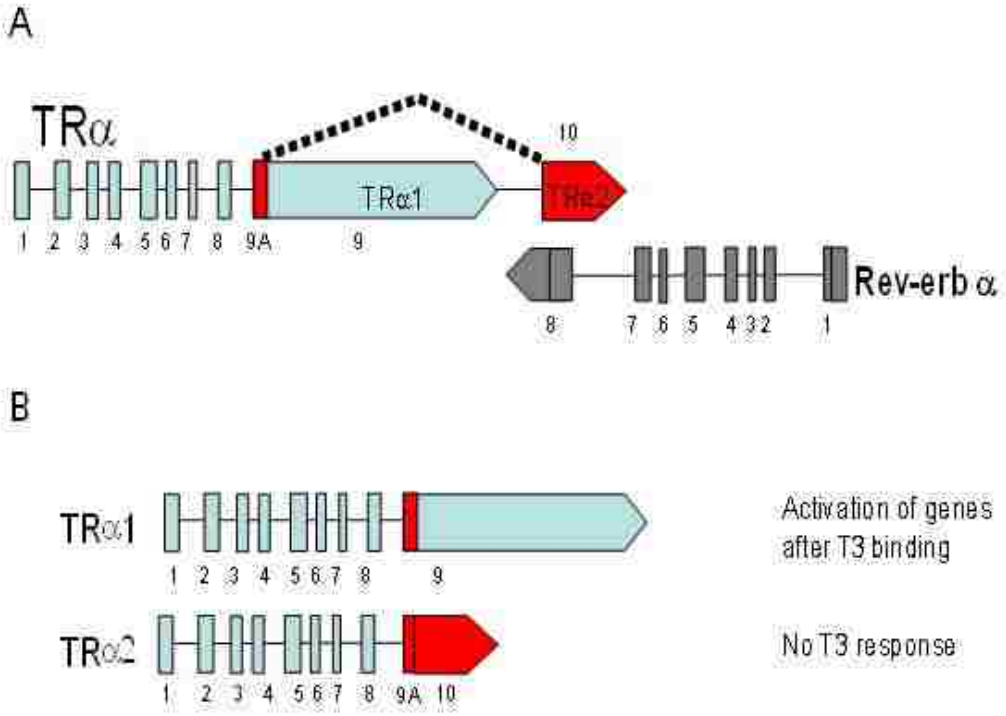


Figure 1 Schematics structure of TRα gene and alternative splicing process

A. In mammals, TRα gene contains 10 exons and produce two RNA transcripts by alternative splicing of pre-mRNA transcripts. A 263 overlapping region is identified between exon10 of TRα gene and exon 8 of Rev-erbα. The line between different colors on exon9 indicate the position of the TRα2 5'ss and downstream exon.
B. Transcripts of two functional antagonistic protein factors. TRα1 contains exon1 to exon9. TRα2 forms truncated transcript with part of exon9 and full length exon10. TRα1 will activate the downstream T3 responsive genes after binding thyroid hormone while TRα2 will not bind T3 and represses genes that TRα1 activates.

TRβ and Rev-erbβ are two other nuclear receptor genes that are closely related to TRα and Rev-erbα. In *Homo sapiens*, TRα and Rev-erbα are located on chromosome 17 while TRβ and Rev-erbβ genes are present on chromosome 3. Unlike TRα and Rev-erbα, TRβ and Rev-erbβ don't share overlap sequence and are separated by approximately 150 kb. Both TRα and TRβ encode for thyroid hormone receptors and mediate transcriptional activity in conjunction with numerous coactivators and corepressors [20, 21, 22]. Our study of TRα alternative splicing will provide insight for general alternative splicing mechanism and the possible regulatory rate of the antisense overlap. It may also contribute to understanding of the regulation of Rev-erbα.

The identification and functional analysis of cis-acting elements in TRα alternative splicing

In eukaryotes, the regulation of alternative splicing requires the participation of RNA sequence (cis-acting elements) and various protein factors (trans-acting factors), which promote or inhibit the assembly of the functional spliceosome. To achieve successful function of sequence motifs, different kinds of RNA-protein and protein-protein interactions are anticipated. [23, 24, 25]. According to their positional and functional difference, cis-acting RNA motifs can be categorized as: exonic splicing enhancer (ESE), exonic splicing silencer (ESS), intronic splicing enhancer (ISE) and intronic splicing silencer (ISS) [26, 27]. Trans-acting protein factor will then bind to these RNA sequences and affect the corresponding splicing activity. Typically, serine-arginine family protein factors (SR proteins) will bind to exonic splicing enhancers and positively affect the splicing efficiency. Many heterogeneous nuclear ribonucleoproteins (hnRNPs) are known

bind to ISS and ESS motifs and inhibit corresponding splicing activity. However, hnRNP F and hnRNP H could contribute to splicing enhancement after binding to certain intronic splicing enhancers [28, 29, 30]. Studies in our laboratory have been focused on the mechanism for regulating alternative splicing of TRα. The role of cis acting elements and polyadenylation on alternative splicing have been investigated in our system (Figure 2).

Production of functional TRα2 mRNA requires cis-acting elements and corresponding tran-acting factors [31]. Previously, Dr. Michelle Hastings from our laboratory identified the specific enhancer SEα2, an 80 nt segment downstream of TRα2 5' splicing site [32]. SEα2 is a highly conserved sequence motif located in exon 9, which is also the final intron of TRα2. Substitution of the purine residues sequence inhibits the efficient splicing of TRα2 *in vitro*. Studies from a series of truncated constructs indicate that deletion of SEα2 induces a sharp decrease of normal TRα2 splicing *in vivo* [32]. The enhancing capability of SEα2 was also examined using heterologous pre-mRNAs. SEα2 segment was added to exon 4 in a truncated drosophila Dsx pre-mRNA [33, 34]. This change increased Dsx splicing efficiency more than threefold comparing to the negative control. SEα2's interaction with trans-acting factors was identified by *in vitro* competition experiments. If SEα2 binds to protein factors necessary for TRα2 splicing, an excess of SEα2 will inhibit TRα2 splicing competitively [32]. In competition experiments, different concentrations of wild-type and mutant competitor RNAs were incubated with TRα2 pre-

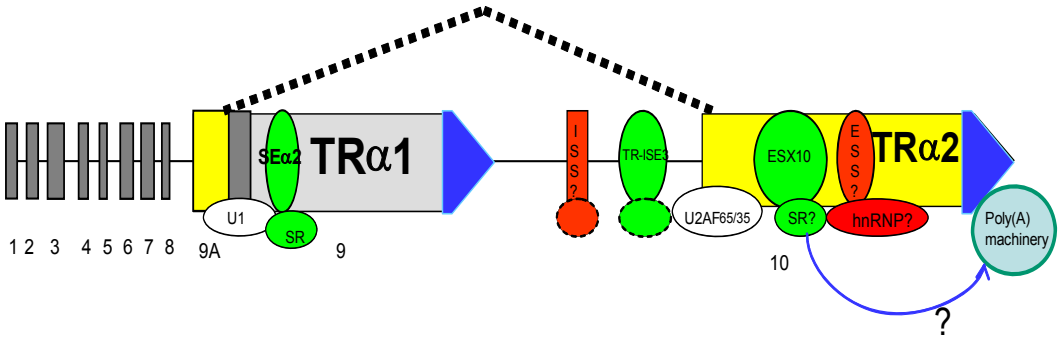


Figure 2. Regulatory mechanisms for TRα alternative splicing.

Alternative of TRα involves different regulatory mechanisms. The interaction of cis acting elements and trans - acting factors is one major category of regulation. The cis acting motifs RNA can be categorized into: exonic splicing enhancer (ESE), exonic splicing silencer (ESS), intronic splicing enhancer (ISE) and intronic splicing silencer (ISS). Splicing silencer elements are shown in red, splicing enhancer elements in green. Both ISS and ESS are hypothetical, while three different kinds of enhancing elements have been characterized. Selected examples of protein known to regulate splicing and polyadenylation are also included: U1 snRNP, U2AF65/35, SR protein, hnRNP proteins. Currently, we have identified SEα2, TR-ISE3 and ESX10 as cis acting elements in our system. Interaction between SEα2 and SR protein factors has been identified[32]. Question mark indicate hypothetical regulation elements shown for illustrative purposes.

mRNA in HeLa nuclear extract for a fixed time period and corresponding TRα2 pre-mRNA splicing efficiency was then evaluated. In contrast to mutant or non-specific RNA competitors, SEα2 derived competitors efficiently inhibit normal TRα2 splicing *in vitro*. Several tran-acting protein factors including SF2/ASF and hnRNPH were identified by protein–RNA crosslinking with radio-labeled SEα2 RNAs [32]. These experiments strongly suggested SEα2 and several associated protein factors are essential for production of TRα2 mRNA. Alternative splicing is often affected by specific sequence elements near the suboptimal 5' or 3'ss [35]. TR-ISE3 (Thyroid Hormone Receptor-intronic splicing enhancer-3'), a different type of cis-acting element was identified and studied by Ms. Valerie Salato in our lab [36, 37]. Multi-method studies with truncated constructs indicates that TR-ISE3 is a 90 nt sequence segment located upstream of 3'splice site of TRα2 (Figure 2). *In vitro* splicing studies suggest that the whole length of TR-ISE3 is also essential for normal TRα2 production. Interestingly, SEα2 and TR-ISE3 seem to have an additive effect on enhancing TRα2 splicing [36, 37], which suggests that they function independently. Such interaction between different cis-acting elements might be critical for alternative splicing.

Another element important for splicing of TRα2 is ESX10 (exonic splicing enhancers on exon10), originally identified by Ms Salato. Because exon10 is the terminal exon for TRα2 and contains the overlapping region of two nuclear receptor proteins, the existence of cis-acting elements might contribute to the regulatory mechanism for transcription of both genes. Initially, a series of overlapping deletions were used to study the necessity of

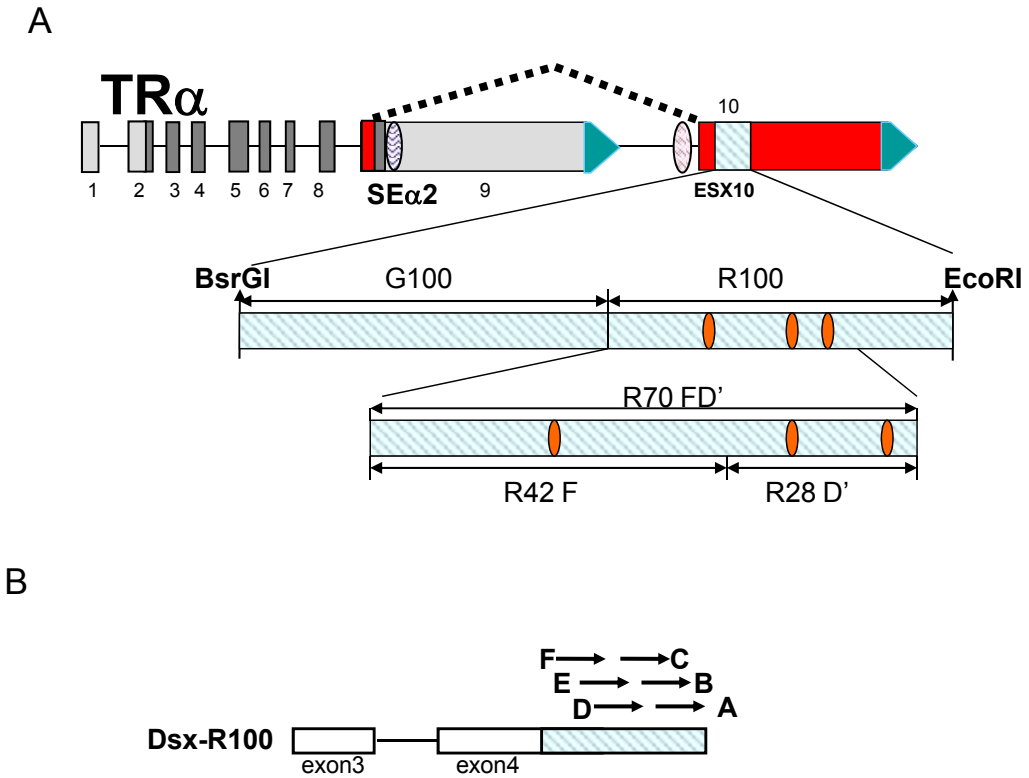


Figure 3 Schematic structure of ESX10 and its subfragment

A. **In order to further characterize ESX10, ESX10 segment (GR fragment) was subdivided equally and named G100 and R100 according to their nearby restriction sites. R100 was divided into 6 consecutive overlapping fragments and named R42αA- R42αF starting from the 3' end. Bioinformatics work predict that three heptamer motifs of R100 might be short ESE targets.**

B. **A truncated drosophila Dsx *in vitro* system was used to study the enhancing capability of specific sequence segments. Without its original enhancer, the pre-mRNA spliced poorly. G100, R100 and other interesting sequences were inserted to the 3' end separately to test enhancing ability. A-F indicate six 42 nt fragments of R100, described above.**

ESX10 and successfully narrowed the size of enhancing fragment to 200 nt [36, 37]. It is located downstream of TRα2 3'ss and named the GR fragment according to the surrounding restriction sites (BsrGI and EcoRI). The enhancing capability of ESX10 has been tested both *in vitro* and *in vivo*. Since Rev-erbβ sequence is 63% identical to the Rev-erbα sequence but shows little enhancing capability in *in vitro* splicing test, it is used in its antisense orientation as a negative control for the study of ESX10. Introduction of ESX10 fragment into Dsx system induced a significant increase of Dsx exon splicing compared to corresponding Rev-erbβ sequence. To characterize better the function of ESX10 in vivo, new constructs have been made with the original ESX10 substituted by corresponding Rev-erbβ sequence. Following transfections of HEK293 cells, total RNA was assayed by realtime PCR. Results indicate that TRα2's splicing was inhibited in the absence of normal ESX10. To further characterize ESX10, the GR fragment was subdivided equally into two fragments, named G100 and R100 according to their adjacent restriction sites. When both G100 and R100 were inserted into Dsx for *in vitro* splicing test, the R100 showed better enhancing capability [36, 37]. To better characterize enhancing segments, R100 was divided into 6 consecutive overlapping fragments and named R42αA-R42αF starting from the 3' end [36, 37]. Each fragment is 42 nt long and overlapping 28nt with neighboring fragments and the enhancing capability of each piece was tested with Dsx *in vitro* system (Figure 3B). All six fragments promoted the splicing activity of Dsx pre-mRNA and their enhancing level were similar, which also indicate the existence of multiple enhancing elements within R100. In order to further characterize critical enhancing motifs, a series of R70 sequence were generated with all possible combination of R42αF (R42βF) and 28nt R42αD' (R42βD') and their

enhancing capabilities were evaluated *in vitro* by Dsx system. The fragments with TRα sequence antisense to Rev-erbα displayed higher activity compared to the antisense Rev-erbβ (R70αFαD' > R70αFβD '> R70βFαD' >> R70βFβD') [36, 37]. This ordering reflects the fractional representation of TRα sequence in each construct. The results from this series of experiments suggested that the authentic TRα sequence is important for the enhancing capability of ESX10 subfragments and we are interested in further studying the enhancing elements within R100.

Alternative polyadenylation of TRα

In addition to splicing, alternative processing of TRα pre-mRNA also involves alternative sites for polyadenylation [39, 40]. Polyadenylation is an important post-transcriptional process for eukaryotic mRNAs and is involved in many processes including facilitating translation, protecting mRNA from degradation in cytoplasm and exporting the mRNA from the nucleus for translation. The process of polyadenylation introduces many adenylate residues to the 3' end of the transcript. In animals, a typical signal associated with polyadenylation is the hexanucleotide sequence AAUAAA located approximately 10-25 nt upstream of the polyadenylation site [40, 41, 42]. Each of the two TRα mRNAs has a unique poly (A) site. Previous work by two students, Ms.Angela Schnell and Ms Elizabeth Mackiel, contribute greatly to further analysis of the effect of polyadenylation on TRα alternative splicing (Figure 2). The simian virus 40 (SV40) late mRNA share a strong poly(A) site (SVL) and its sequence is used to represent the effect of strong polyadenylation.

A series of constructs have been made with SVL poly (A) site introduced to either TRα1 or TRα2 poly(A) site as a 175 nt, BamHI fragment. RNase protection assays (RPAs) and realtime PCR were used to study TRα1 and TRα2 expression level. These studies showed that SVL poly (A) signal substituted in TRα1 exon will aggressively compete with normal TRα2 splicing and result in almost exclusive production of TRα1. When the downstream TRα2 poly(A) site was replaced with the strong SV40 signal, the balance between TRα1 and TRα2 mRNA expression was also altered, resulting in increasing amounts of TRα2 splicing [S. Munroe and A. Schnell unpublished]. Unlike the exclusive TRα1 production result from TRα1 poly (A) substitution, the downstream SV40 signal will enhance TRα2 splicing but will not compete out TRα1 mRNA splicing. It appears that both upstream and downstream poly (A) sites are critical for maintaining TRα1/ TRα2 ratio. Introduction of a stronger TRα2 5' splice site with a single nucleotide substitution (+5C/G) counteracts the effect of strong poly (A) signal of TRα1 and restore the normal balance. Therefore, the strength of splice and polyadenylation sites also affects the balance between splicing and polyadenylation. Since both cis-acting splicing regulatory elements and polyadenylation are important for TRα alternative splicing it is important to study the interactions between polyadenylation and ESX10 and investigate their effects on the balance between TRα1 and TRα2. Using TRα system as model, we hope to explore the underlying regulatory mechanism of alternative processing and increase our understanding of the complexity of gene expression.

B. Specific aims

As stated in the above introduction, the major goal of my study is to further identify and investigate cis-acting elements in alternative splicing of α-thyroid hormone (T3) receptor gene. Since polyadenylation also affects this process, the effect of strong poly(A) signal on TRα alternative splicing has also been studied. The specific aims of my study are:

1. To characterize three ESE motifs within R100 region of ESX10 and further study the functional roles of ESX10 subfragments. These studies on ESE motifs are carried out by substituting ESE motifs in exon 10 with corresponding homologous sequence copied from the antisense strand of the Rev-erbβ gene sequence. The enhancing effects of the chimeric fragments are tested by *in vitro* studies with Dsx system. *In vitro* competition experiments are also used to characterize the functional role of ESE motifs in TRα2 alternative splicing. To further study the requirements for ESX10 activity, substitution of either half of ESX10 (GR) has been examined in transfection experiments using a truncated TRα minigene plasmid.

2. To investigate the effect of polyadenylation sites in regulating TRα alternative splicing. Previous work indicates that inactivation of ESX10 blocks the normal TRα2 splicing. This effect of substituting a strong poly(A) site at 3' end of TRα2 will be studied with respect to the activation of cryptic splicing within G100 and R100 segments. Interestingly, the same strong poly(A) site enhances TRα2 splicing fidelity by simultaneously reducing TRα3 expression, a shorter variant of TRα2 with a 3'ss within G100

Chapter II. Characterization of Bioinformatically Defined ESE Motifs within ESX10 Enhancer

A. Introduction

In the Introduction we described the identification of ESX10 and the analysis of its enhancing capability through *in vitro* splicing. In order to analyze comprehensively the presence of typical ESE motifs within ESX10, we used bioinformatics approaches to study its sequence and provide insights for further studies. Through bioinformatics sequence analysis with Rescue-ESE program [43], three possible ESE motifs were identified within the R100 fragment of ESX10. To further analyze their functional roles, a series of *in vitro* experiments are used to study these ESE motifs. In this section I describe the results from the bioinformatics studies, following *in vitro* splicing and competition experiments to characterize the role of these ESE motifs in ESX10's enhancing effect.

B. Further study of possible ESE motifs within R70 fragment

Since many ESE elements are 6-7 nucleotides segment that are sufficient for binding specific SR proteins, several well-known programs including Rescue-ESE and FAS-ESS, ESE-finder and Splicing rainbow are applied to analyze the ESX10 sequence and predict possible ESE motifs. [43, 45, 46, 47]. By analyzing R100 sequence with Rescue-ESE program [43], three heptamers have been identified and further analyzed. Although even more clusters of Rescue hexamers are detected besides R100 in exon 10 (Figure 4), these hexamers in R100 are intriguing since they overlap five of six previously tested 42nt

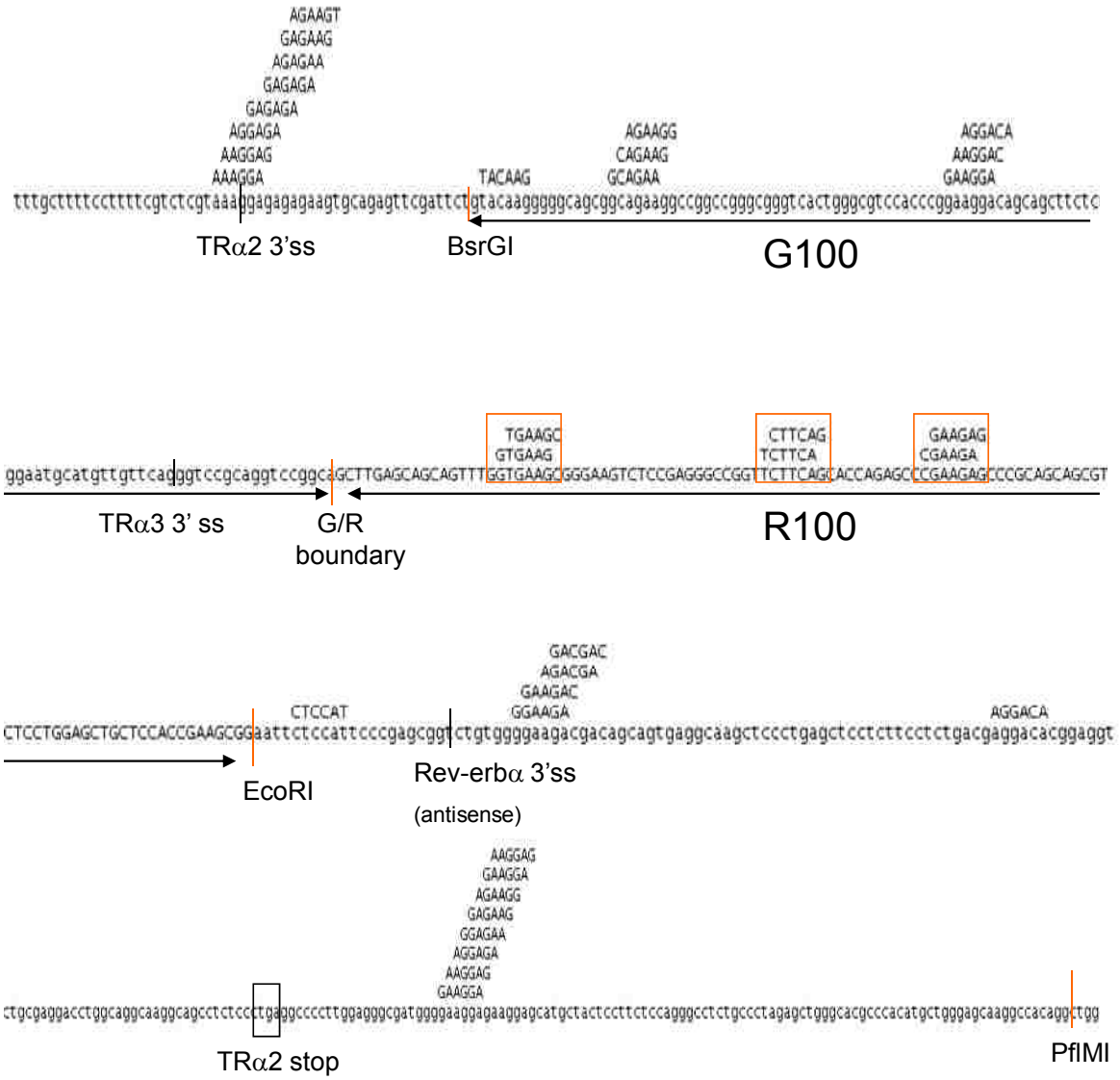


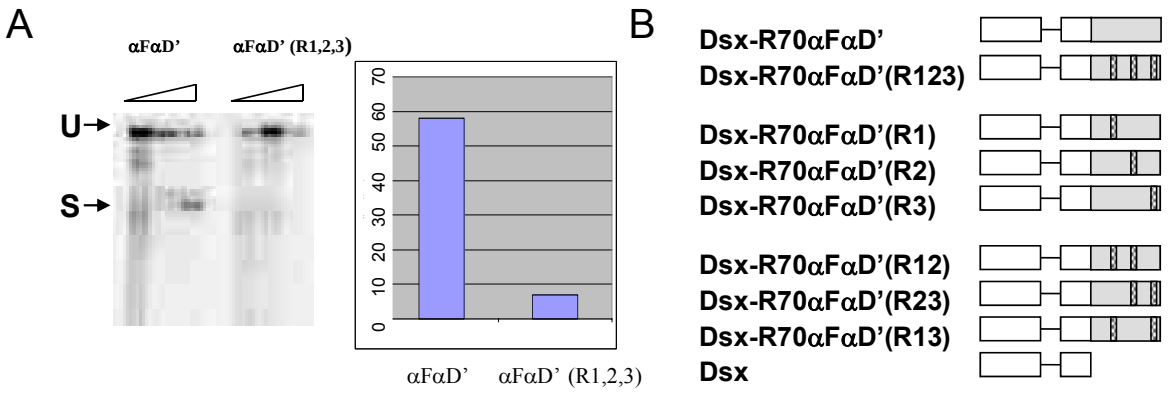
Figure 4.Sequence information about ESE motifs on exon10.

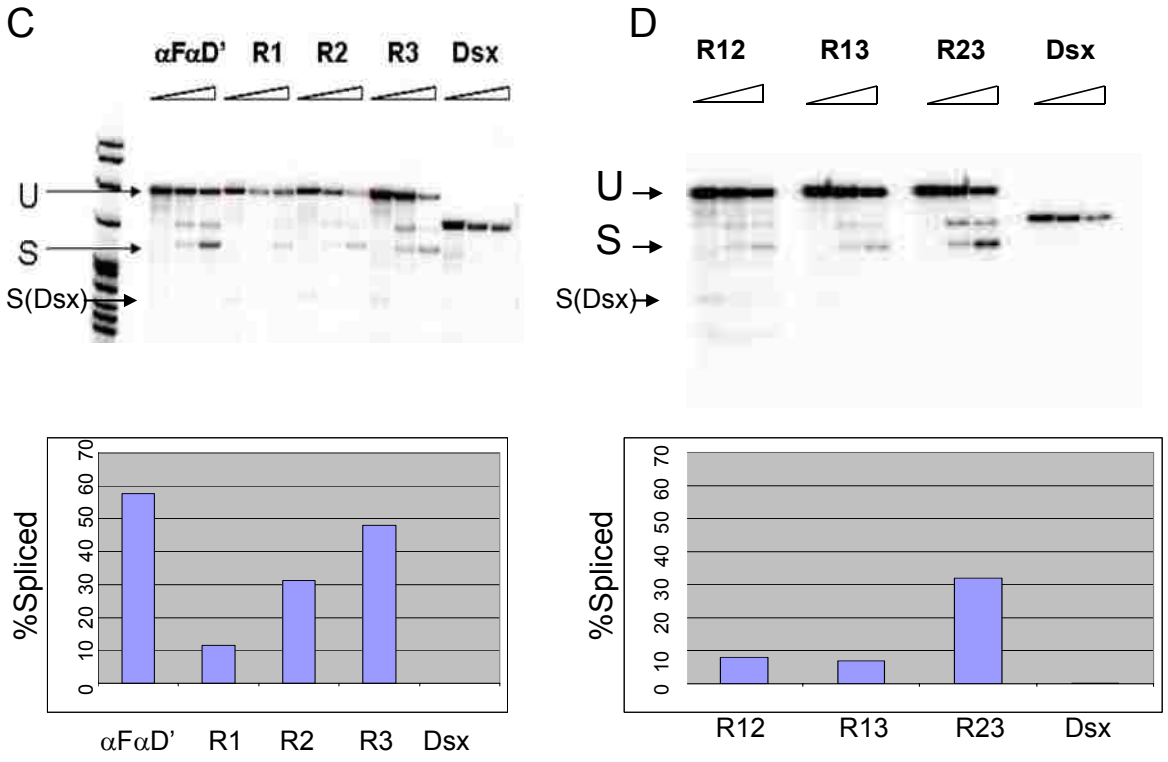
ESX10 Sequence analysis with Rescue-ESE program. All predicted ESE hexamers within ESX10 have been marked above the sequence. The box circled out are selected possible ESE elements within R100 region. Each one of them is 7 nt long and comprised of two overlapping ESE hexamers. BsrGI and EcoRI sites define the end of GR fragment. EcoRI and the first PflMI site define the ends of the RM fragment.

subfragments (R42αB-R42αF)[36]. Each heptamer contains two overlapping predicted ESE hexamers and they are named Rescue1 (R1) Rescue 2(R2) and Rescue3 (R3) depending on positional difference (Figure 4). After predicting these ESE elements' position, various *in vitro* experiments are performed to directly test their effect on ESX10 enhancing capability.

1. Substitution of all three ESE motifs with Rev-erbβ sequence cause a sharp decrease of the enhancing capability of ESX10 subfragment.

In order to determine whether these ESE segments are essential for ESX10's activity, all three motifs were substituted with the corresponding Rev-erbβ sequence. In the Rev-erbβ sequence, R1 has 3 nucleotide substitution, R2 has 2 nt changed and R3 has 3 different nucleotide compared to the authentic TRα sequence (Figure 5E). Overall, eight nucleotide substitutions were introduced into R70αFαD' (R123) and the modified fragment was inserted into truncated Dsx exons to test its enhancing capability. After adding chimeric fragment into Dsx system, the splicing activity (1.4%) dropped dramatically compared to addition of the original R70αFαD' sequence (60%). Thus, the mutation of all three ESE motifs results a loss of 95% of the splicing activity (Fig 5A). These results suggest that the authentic sequences of these ESE heptamers are essential for maintaining the enhancing activity of ESX10 subfragments *in vitro*.


A
αFαD'
αFαD' (R1,2,3)
U
S
0 10 20 30 40 50 60 70
αFαD' αFαD' (R1,2,3)
B
Dsx-R70αFαD'
Dsx-R70αFαD'(R123)
Dsx-R70αFαD'(R1)
Dsx-R70αFαD'(R2)
Dsx-R70αFαD'(R3)
Dsx-R70αFαD'(R12)
Dsx-R70αFαD'(R23)
Dsx-R70αFαD'(R13)
Dsx


C
αFαD' R1 R2 R3 Dsx
U
S
S(Dsx)
%Spliced
0 10 20 30 40 50 60 70
αFαD' R1 R2 R3 Dsx
D
R12 R13 R23 Dsx
U
S
S(Dsx)
%Spliced
0 10 20 30 40 50 60 70
R12 R13 R23 Dsx

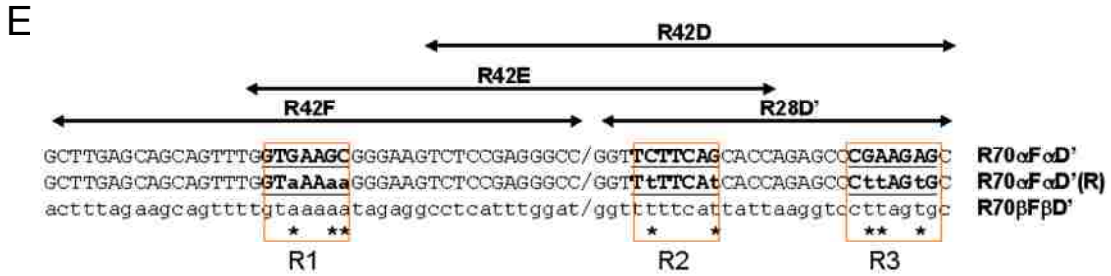

E
R42D
R42E
R42F
R28D'
GCTTGAGCAGCAGTTTGGTGAAGCGGGAAGTCTCCGAGGGCC/GGTTCTTCAGCACCAGAGCCCGAAGAGC R70αFαD'
GCTTGAGCAGCAGTTTGGTaAAaaGGGAAGTCTCCGAGGGCC/GGTTtTTCAtCACCAGAGCCCttAGtGC R70αFαD'(R)
actttagaagcagttttgtaaaaatagaggcctcatttggat/ggtttttcattattaaggtccttagtgc R70βFβD'
R1
R2
R3

Figure 5. Substitution of nucleotides in ESE motifs reduce the enhancing capability of R70 subfragment

A. **In vitro splicing results showing Dsx pre-mRNA with inserted R70 fragments. The αFαD' represent R70 with all TRα sequence and αFαD'(R1,2,3) indicate the R70 fragment with all ESE motifs substituted with Rev-erb β sequence. Calculated splicing percentage are indicated on the left bar graph.**

B. **Schematic figures about constructing substrates that are used in in vitro splicing. Structure of Dsx pre-mRNA with inserted αFαD' R70 fragment is on the top while the negative control with truncated Dsx sequence is one the bottom.**

C. **In vitro splicing results showing Dsx pre-mRNA with inserted R70 fragments that contains substitution of each single ESE motifs with Rev-erbβ sequence. The results from left to right indicate the splicing reactions of αFαD', αFαD'(R1), αFαD'(R2), αFαD'(R3). The calculated splicing percentage is shown on the bottom.**

D. **In vitro splicing results showing Dsx pre-mRNA with inserted R70 fragments that contains substitution of double ESE motifs with Reverbβ sequence. The results from left to right indicate the splicing reactions of αFαD'(R12), αFαD'(R13), αFαD'(R23) and Dsx. The calculated splicing percentage is shown on the bottom. (cooperation with Dr.Munroe)**

E. **Sequence of R70αFD', R70FD'(R123) and R70βFβD' fragments. In this figure, α and β stands for TRα and Rev-erbβ separately and the sequence of TRα are in capital letters while Rev-erbβ sequence are shown in lower case. The difference between α and β sequence are marked with asterisks. Arrows indicate the subdivided ESX10 fragments R42F, R42E and R42D. Boxed sequences are three ESE hepatmers R1,R2 and R3.**

2. Substitution of single or double ESE motifs with Rev-erbβ sequence reduces enhancing capability of ESX10 subfragments *in vitro*.

After observing that the substitution of all three ESE motifs caused a large decrease in R70 activity, we further studied the functional characteristics of each ESE element and the interactive relationships between them. Similar to the strategy used to investigate the effect of all three ESE heptamers, the substitution of each ESE motif was performed independently with corresponding Rev-erbβ sequence and the mutated R70 sequences were added to truncated Dsx exons to examine their effects (Figure 5B). The *in vitro* splicing results showed that after introducing mutation to single ESE motif, the enhancing activity of chimeric fragments decreased for every case. Among three ESE motifs, the substitution of three nucleotides in R1 induces the most significant negative effect on R70 enhancing capability, causing about 80% reduction on splicing activity compared to the intact R70 fragment (Figure 5C; splicing percentage: 11% in R1 compared to 58% in αFαD'). Compared to the sharp drop in Dsx pre-mRNA splicing after R1 substitution, the substitution of either R2 or R3 individually had slight effect on splicing activity (Figure 5C: splicing percentage: 31% in R2 and 48% in R3 compared to 58% in αFαD'). From these results, we concluded that mutation of single ESE motifs could also attribute to the enhancing effect of ESX10 subfragments and R1 motif might be the critical for ESX10's overall enhancing function.

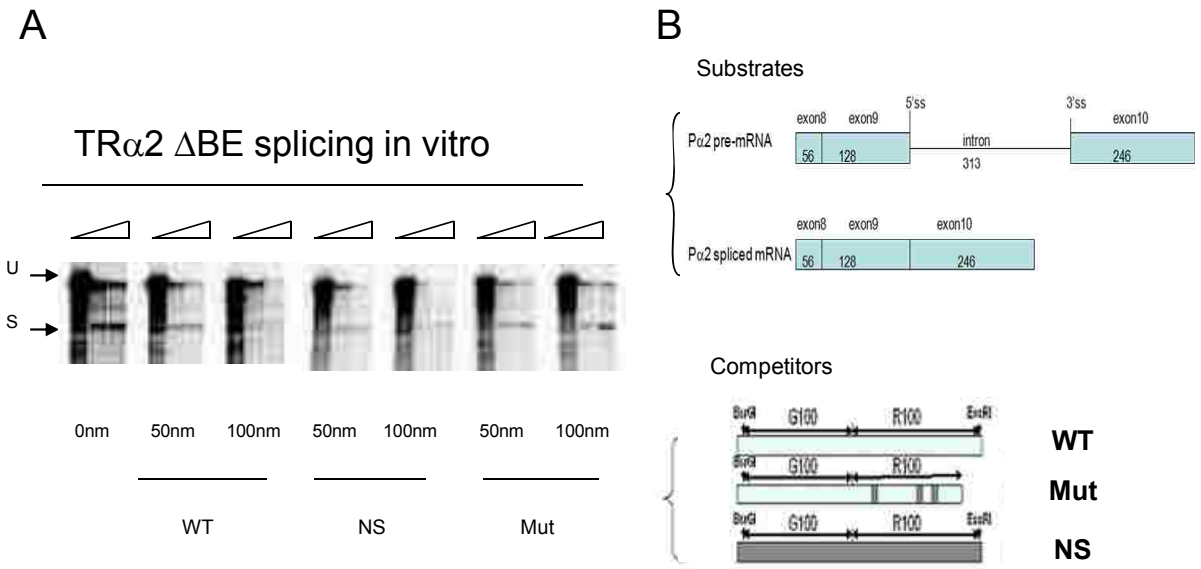
The successful splicing inhibition after substitution R1 ESE motif raised another interesting question: is R1 the only contributor for ESX10's function or are the contributions of other ESE elements (such as the R2 and R3 motifs) also required to

achieve normal enhancing effect? To answer this question, we substituted every two out of three ESE motifs and tested their effects with similar methods to investigate the interaction between these elements (Figure 5B). The *in vitro* splicing results strongly suggested that the ESE motifs would interact with each other and work cooperatively. Mutation of R2 and R3 motifs alone only induced slight inhibitory effect on splicing activity compared to R123 fragment (Figure 5D; splicing percentage: 32% in R23 compared to 65% in αFαD' control; please see reference 36 for respective αFαD' data). And their mutations combined with R1 (R12 and R13) were much more effective in inhibiting R70 enhancing activity (Figure 5D; splicing percentage: 8% in R12 and 7% in R13 compared to 65% in αFαD' control; please see reference 36 for respective αFαD' data). Compared to αFαD' control, he inhibitory effects from the double mutations with R1 (R12 and R13) are slight lower than the substitution of all three motifs (R123) but higher than substituting R1 alone (R1) [36]. Considering the close positional distance and functional relationship between R2 and R3, it might be more reasonable to regard R2 and R3 as another functional element (R2/R3) for ESX10. These results confirmed the critical role for R1 heptamer on R70's enhancing capability and suggested that multiple enhancing elements within R100 might cooperate with each other to promote TRα2 splicing.

3. Functional analysis of the ESE motifs by competition experiments.

Since many ESE motifs commonly bind members of the SR protein family, it was possible that the trans-acting proteins within nuclear extract would interact with these predicted ESE motifs. As a first step towards characterizing the involved protein factors,

in vitro competition experiments were performed. If the ESE elements binds to the protein factors present in nuclear extracts and these interactions are essential for the enhancing capability, an excess of normal R70 RNA fragment with original ESE sequence would competitively occupy necessary protein factors and inhibit the normal splicing of TRα2 competitively. In contrast, an excess of chimerical fragment with mutated ESE would not affect TRα2 splicing. In competition experiments, splicing reaction containing radio-labeled truncated TRα2 were incubated with different competitors RNAs (Figure 6A). The substrate pre-mRNA is transcribed from pα2 ΔBE construct which contains partial TRα exon9 and exon10. Originally, the RNA substrate is 743 nt long and a final 430 nt product will be detected after intron removal from *in vitro* splicing. The competitor RNA fragments include wildtype sequence (competitor WT: a 220 nt RNA fragment with wildtype ESX10 sequence), mutated sequence (competitor Mut: a 170 nt RNA fragment with ESX10 that has 3 ESE motifs substituted with Rev-erbβ sequence) and corresponding Rev-erbβ sequence (competitor NS: a 220 nt RNA fragment with Non-specific Rev-erbβ sequence) (Figure 6B). Besides, all competitors have been applied to the pα2 ΔBE splicing with different concentrations. According to the preliminary data, addition of 100nm competitor WT (original ESX10 sequence) results in a 20% decrease on the splicing activity (Figure 6C). This result reflects specific interaction between ESX10 and certain trans-acting factors present in the nuclear extract because the mutated competitor Mut only have a slight inhibitory effect on splicing. Comparatively, addition of the negative competitor NS (Rev-erb β sequence) didn't induce obvious competition effect either. These results suggest that the trans-acting



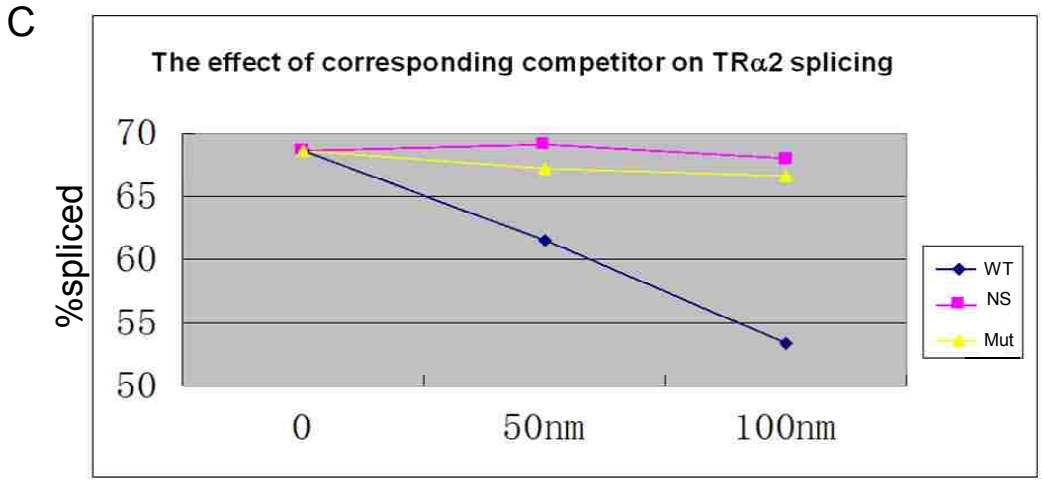


Figure 6. Inhibiting of TRα2 splicing with excess competitor RNA

A: **The pα2 ΔBE pre-mRNA was incubated with 0, 50, or 100 nM of the labeled competitor WT(wildtype: ESX10 sequence), competitor NS (Non-specific Rev-erb β sequence) and competitor Mut(ESX10 sequence with mutated ESE motifs). All reactions were incubated for 90 minutes and the samples were collected at 0 min, 45 min, 90 min.**

B: **The schematic structure of pα2 ΔBE pre-mRNA and competitor RNAs used in the experiments**

C. **Percentage of Splicing for pα2 ΔBE in vitro splicing with different competitors.**

factors in Hela nuclear extracts are required for ESX10 activity and the previous identified ESEs within this region might be critical in this cis-trans interaction.

C. Discussion

Regulation of alternative splicing involves the participation of various mechanisms and one major contributor is the existence of different cis-acting elements and following cis-trans interactions. In this chapter, my studies focused on the further characterization of three ESE heptamers elements located within the 3' half of ESX10. These ESE motifs were original identified through detailed sequence analysis of ESX10 with bioinformatics methods and their effects on the enhancing capability of ESX10's subfragments are investigated by *in vitro* splicing. Afterwards, competition experiments are applied to further explore the possible cis-trans interaction within ESX10 region. My studies indicate that the mutation of eight nucleotides within these ESEs dramatically reduced the enhancer activity of ESX10 subfragments *in vitro* and also weakened the possible cis-trans interaction.

The studies related to Exonic splicing enhancer motifs

In order to further study these ESE motifs within ESX10 region, the ESE sequences were investigated by replacing single or all ESE motifs with corresponding Rev-erbβ sequence and the enhancing capability of chimeric fragments are tested *in vitro*. Although my experiments indicate that the mutation of ESE heptamers induced significant negative effect on the enhancer activity of ESX10 subfragment, it is also possible that the effect is not from the elimination of original enhancer element but from the introduction of new

silencer. FAS-ESS, a bioinformatics program that can predict the existences of ESS were applied to test this hypothesis. After substituting three heptamers with Rev-erbβ sequence, a reduction on number of possible ESE motifs and increase of potential ESS candidates is predicted. In order to further study whether the inhibitory effect is from the disturbance of original TRα2 sequence in heptamer motifs or from possible introduced ESS motifs, we used other sequences combination besides the Rev-erbβ substitution to test the enhancer activity of ESX10 subfragments. If other changes within these ESE regions could reduce the splicing activity, the uniqueness of these ESEs could be confirmed. Because R1 showed most significant effect among these three ESE motifs, we chose this ESE heptamer and changed its sequence by random scrambling the order of the nucleotides to disrupt the original ESE sequence (Generated with primers 774R and 775R). These scrambled motifs were introduced into Dsx system and tested similar as other chimeric segments. Scrambled R1 sequence within the αR42F fragment showed reduced enhancing effect than the original TRα sequence (for example, 35% splicing αR42F compared to 23.2% splicing scrambled R1), although to a less extent than the 3nt Rev-erbβ substitution(1.4%). The results from scrambled experiments suggest that the authentic sequence of these ESE hepatmers might be critical for their enhancing effect and disruption of the nucleotides constitution would negatively affect their function. However, these *in vitro* splicing reactions showed high levels % input degradation(results not shown) so further work is required to quantify the induced negative effect of these substitutions.

Chapter III. Functional Analysis of ESX10 and its Subfragments (R100, G100)

A. Introduction

In Chapter II of this thesis, I described the positive effect of ESE elements located within R100 region of ESX10. My goal in this chapter is to further characterize the importance of ESX10 and its subfragments. Since both halves of ESX10 (G100 and R100) are highly conserved, constructs with specific substitutions were created and used to further analyze their effect on TRα2 splicing *in vivo* following transient transfection of HEK293 cells. In this section, I describe the results from RNase Protection Assays(RPAs) and realtime PCR experiments to characterize the functional role of ESX10 and its subfragmnets.

B. Study of ESX10, G100, R100 and the related cryptic splicing

1. Analyzing the effect of ESX10 and its subfragments by RNase Protection Assay

In order to investigate the functional importance of ESX10 and its subfragments on TRα2 alternative splicing, a series of TRα minigene plasmids with mutated ESX10 were generated for expression *in vivo*. After equally dividing ESX10 into two segments designated G100 and R100 according to their nearby restriction sites, plasmids that had substitution of each half of ESX10 with Rev-erbβ sequence were created for further studies: pErbAmβG100 (G100) and pErbAm-βR100 (R100). Another construct with the whole ESX10 substituted by corresponding antisense Rev-erbβ sequence was generated by Ms Schnell previously: pErbAm-βGR200 (GR200). After transfecting these three constructs and other plasmids (the vector pRC/CMV as a negative control and

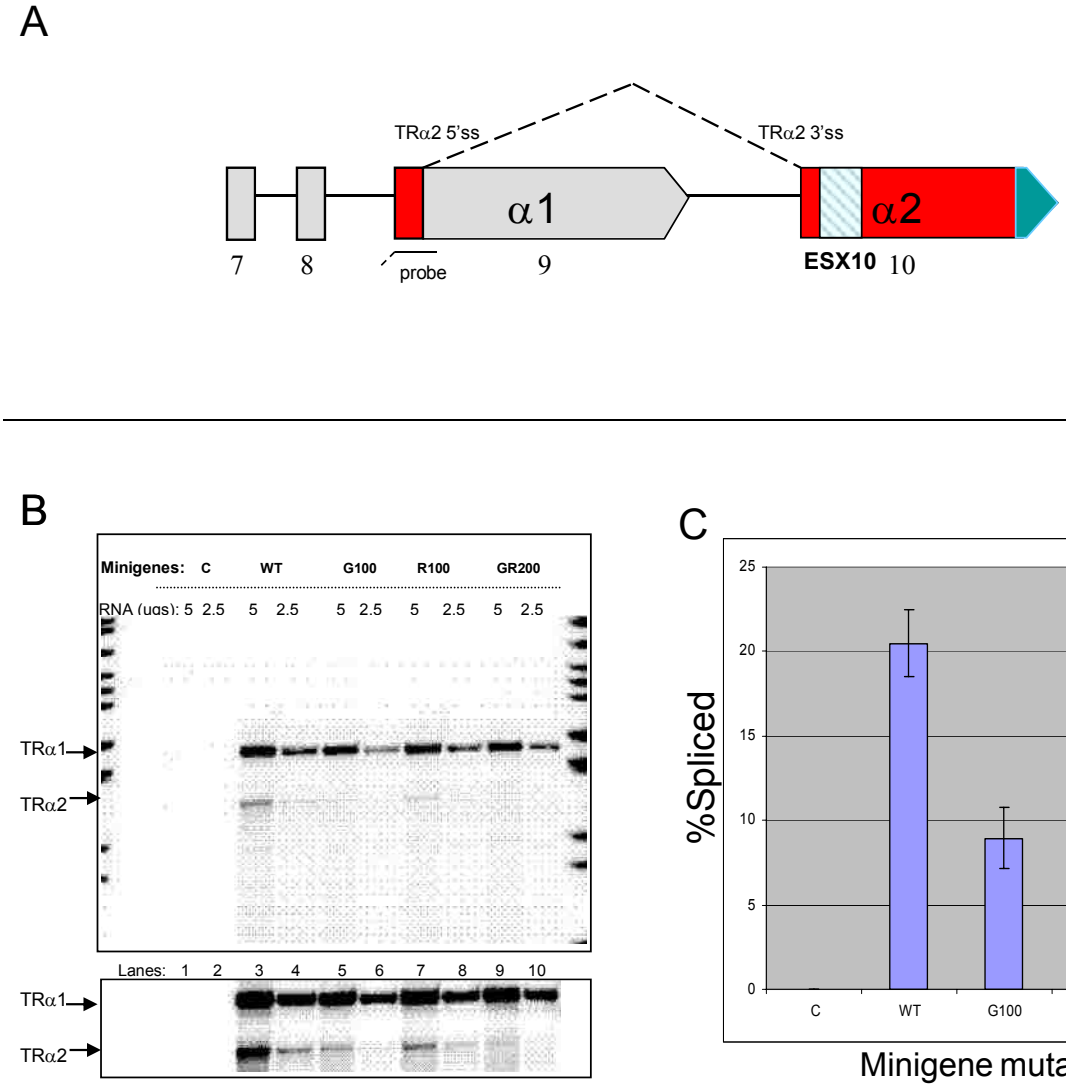


Figure 7. Analyzing the effect of ESX10 or its subfragments by RNase Protection Assay

A. The schematic structure of minigene plasmid we used in transfection. 293 HEK cells were tranfected and total RNA were collected and assayed afterwards.

B. The probe used in the RPA overlaps 5'ss of TRα2, which indicate the expression of both TRα1 and TRα2 spliced mRNA. The arrows indicate the corresponding bands of either product. The tested samples are different constructs with substituted ESX10 or its subfragments. The bottom panel is a high contrast version of the same RPA assay. The abbreviations used in RPA and realtime PCR calculations are: C- Control Vector, WT- wildtype control (pCMVerbAα). G100- Substituting G100 of ESX10 with Rev-erb β sequence (pErbAm Revβ G100). R100- Substituting R100 of ESX10 with Rev-erb b sequence (pErbAm Revβ G100). GR200 - Substituting ESX10 with Rev-erb b sequence (pErbAm Revβ G100).

C. The bar graph on the indicates the calculated TRα2 splicing percentage. The data used here were averaged from at least three independent assays. RNAs are from two independent transfections assayed in duplicate with different amount of input RNA.

pErbA as a wildtype control) into HEK293 cells, total RNA was extracted and used as substrate for RPA assays. RNAase Protection Assay was performed in order to compare the effect of the minigene constructs on TRα2 splicing. The radio-labeled probe used in RPA overlapped the TRα2 5' splice site sequence in exon 9A (Figure 7A). Since TRα1 and TRα2 mRNA contain the shared exon9A sequence, the 5'ss probe will hybridize and protect both products. In the assay, two different RNA concentrations (5 ug and 2.5 ug) were used for each minigene and two independent experiments were performed. The relative splicing levels of both TRα1 and TRα2 were calculated from two separate experiments. The resulting data show that in the minigene without original ESX10 sequence (GR200), the splicing of TRα2 reduced greatly (2.8% splicing) compared to wildtype control WT (20.5%) (Figure 7B and 7C). Substitution of either ESX10 subfragment (G100 or R100) also caused reduction of TRα2 compared to the wildtype control (G100:9.0% and R100:11.5%), but their decreasing folds were lower than substitution of the full ESX10 sequence (GR200) (Figure 7B and 7C). These results suggest that ESX10 is functional important for TRα2 splicing *in vivo* and substitution of ESX10 and its subfragments with Rev-erbβ sequence will also reduce normal TRα2 splicing.

2. Analyzing the effect of ESX10 or its subfragments by realtime PCR assay

Since the above RPAs would measure all splicing events at TRα2 5'ss, we use realtime PCR to confirm the specific inhibition of TRα2 splicing. With primers specific for TRα1 and TRα2, realtime PCR assays were performed to further characterize the functional importance of ESX10 and its subfragments (Figure 8A). From realtime PCR results, we

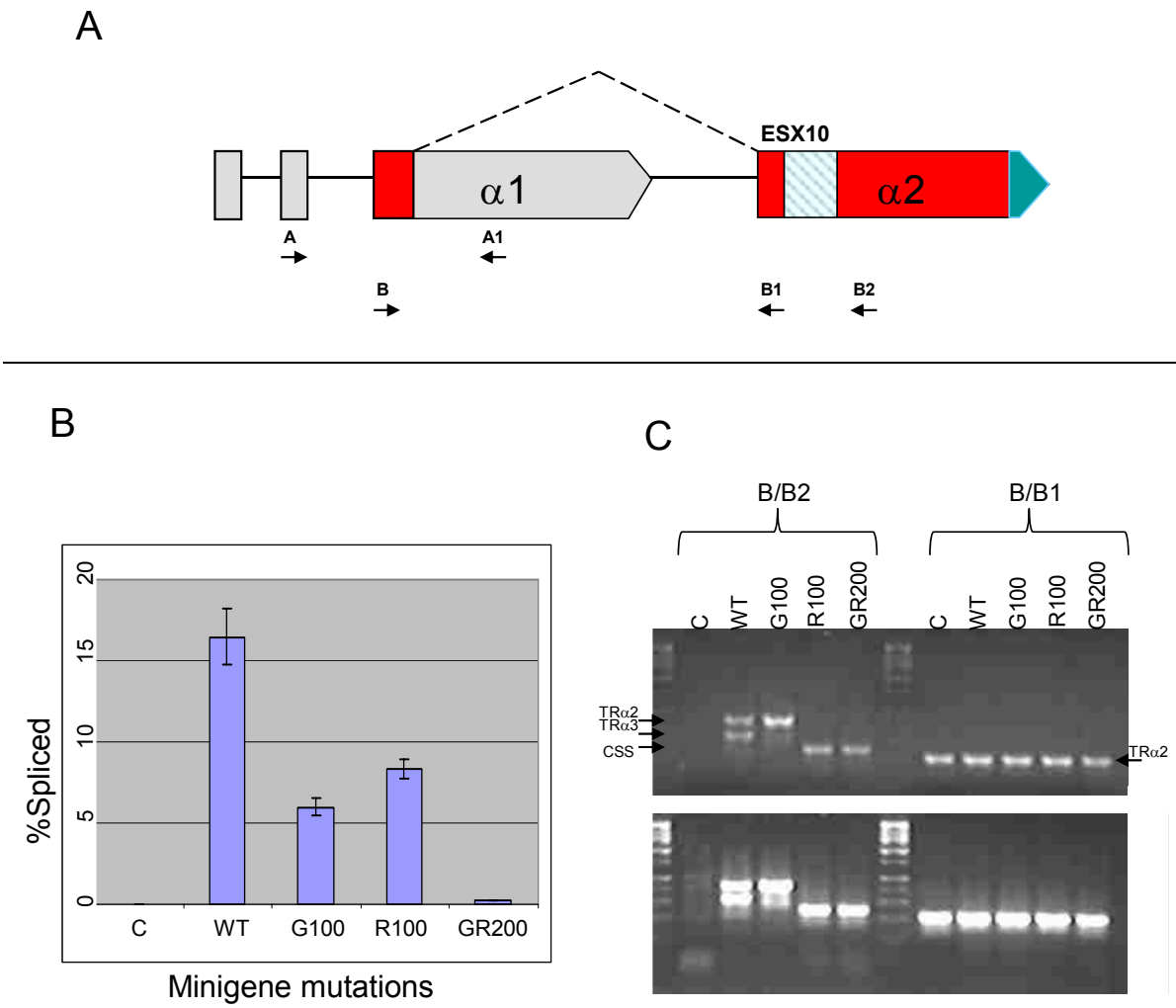


Figure 8. Analyzing the effect of substitution of ESX10 or its subfragments with Rev-erbβ by Realtime PCR methods

A. The schematic structure of minigene plasmids we used for transfection. After collection of total RNA, Reverse-Transcription PCR were used to generate cDNAs as templates for realtime PCR. Various primer pairs were used to analyze the expression level of TRα1 and TRα2. In the figure, A and B stand for primer set designed for different purposes. A- 936F, A1-937R, A2-1499b. B- 900F, B1-1467R, B2- 905R. The sequence of primers are included in Appendix- E.

B. The calculated splicing percentage of TRα2 based on Realtime PCR data with primer set A/A1 and B/B1. The data used were averaged from two independent Realtime reactions and each reaction had three parallel assays, using primers pairs A/A1 for TRα1 and B/B1 for TRα2. % splicing is calculated from the difference between C_T values for TRα1(A/A1) and TRα2(B/B1) realtime PCR as described in Methods (Chapter VI). These percentages assume negligible constrictions from alternatively spliced form including TRa3 and cryptic splicing within βGR200 sequence.

C. Selected samples were assayed with 2% Agarose gel to test their matched size. The templates and primer sets were indicated above the figure. The lower panel is a high contrast version of these bands. CSS stands for products from cryptic splicing sites.

confirmed our findings that substitution of both subfragments showed decreasing TRα2 splicing (Figure 8B). However, PCR analysis of minigene expression also revealed differences. Using the primer set B/B1 specific for TRα2 detection, all the minigene constructs expressed detectable levels of TRα2 mRNA as seen by electrophoretic analysis (Figure 8C). However, with primer set B/B2, multiple spliced products were observed and difference in expression patterns became apparent (Figure 8C). In the minigene with wildtype, a shorted product was evident, which matches the size of TRα3, a previous reported TRα isoform[13,14]. The subsequent sequencing of the additional product confirmed the existence of TRα3 in our system and further studies about TRα3 will be described in Chapter IV. With primer B/B2, G100 showed similar pattern as wildtype control (WT), where TRα2 and a band with similar size to TRα3 were detected. However, substitution of G100 with Rev-erbβ sequence eliminate key residues within the TRα3 3'ss (CAG/C →TTT/A). After analyzing the sequence results of the respective bands, it is clear that the product with TRα3 size is not spliced from the authentic 3' TRα3 splice site but from a cryptic splice site in the substituted Rev-erbβ G100 sequence, which is 2nt downstream the position of authentic TRα3 splicing site (Table I). On the other hand, R100 showed a similar banding pattern as GR200, where a different product from cryptic splicing was predicted. The realtime PCR results indicate that both G100 and R100 have enhancing effect and substitution of each subfrgament would induce decrease on TRα2 splicing. However, the electrophoretic analysis with primers B/B2 indicated that the substitution of R100 induced greater inhibitory effect on TRα2 splicing (in Figure 8C, G100 showed normal TRα2 band while R100 didn't), which suggest that R100 is functionally more important than G100.

Table I. Sequence for 3'splice sites ofTRα2 and TRα3

	position	sequence
Consensus	N/A	YYYYYYYYNYAG/GNN
TRα2 3'ss	+1	CGTCTCGTAAAG/GAG
TRα3 3'ss	+117	CATGTTGTTCAG/GGU
TRα2 3'ss CSS(R100)	+194	TTCATTATTAAG/GUC
TRα3 3'ss CSS(G100)	+119	TGTTGTTTAAAG/AUC

In this table, R stands for purine nucleotides, N for random nucleotides. 3'ss and CSS respectively represents 3' splicing site and cryptic splicing sites.

After isolating the additional products from R100, sequencing analysis confirmed the existence of cryptic splicing product and its splicing site (Table I). The cryptic splicing site located 197 nt downstream of TRα2 3'ss and 65 nt upstream of EcoRI site of GR fragment, which belonged to the R100 region of substituted ESX10. We also identified this cryptic splicing product in GR200 and more studies would be introduced in Chapter IV.

Studying the expression level of cryptic splicing through realtime PCR

Based on the data from RPA assays and realtime assays, we concluded that substitution of ESX10 with Rev-erbβ sequence (GR200) caused dramatically decreased TRα2 splicing and substitution of either ESX10 subfragment (G100, R100) reduced TRα2 splicing moderately compared to complete ESX10 substitution (Figure 9B). However, the substitution of antisense Rev-erbβ sequence in ESX10 also introduced a cryptic alternative 3'splice site. Therefore, the observation of cryptic splicing in G100, R100 and GR200 raises the question whether the absence of ESX10 negatively affects the overall splicing level or simply reduces TRα2 splicing accuracy? To answer the question, we used realtime PCR methods with different primer sets to measure the expression level of cryptic splicing.

In the experiments, primer set A/A2 was used to estimate the total expression of TRα mRNAs. The percent splicing of TRα1, TRα2 and additional products are calculated based on the realtime threshold value (C_T) from different primer sets: A/A1 for TRα1 and B/B1 for TRα2. B/B2 is designed to estimate the expression level of other splicing

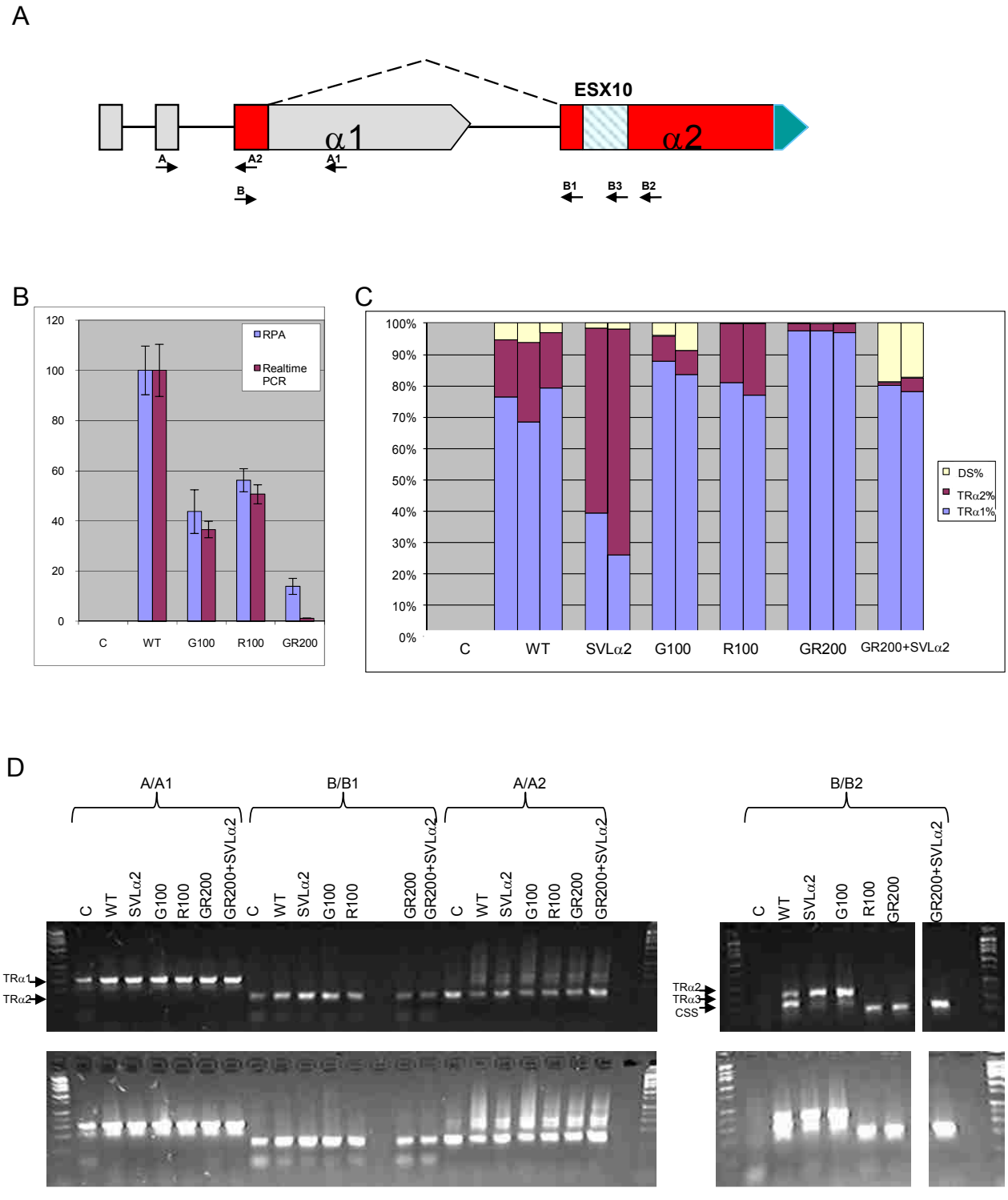

A
ESX10
α1
α2
A
A2
A1
B
B1
B3
B2
B
RPA
Realtime PCR
C
WT
G100
R100
GR200
C
DS%
TRα2%
TRα1%
WT
SVLα2
G100
R100
GR200
GR200+SVLα2
D
A/A1
B/B1
A/A2
B/B2
TRα1
TRα2
TRα2
TRα3
CSS

Figure 9 Comparing the splicing level of different components among minigenes

A. The schematic structure of minigene plasmids we used for transfection. Various primer pairs were used to analyze the expression level of TRα1 and TRα2. A- 936F, A1-937R, A2-1499b. B- 900F, B1-1467R, B2- 1495R, B2- 905R.(See Appendix- E for details)

B. Comparison of the calculated TRα2 splicing percentage between RNase Protection Assays and Realtime PCR. TRα2 splicing in different minigene construct was compared to the wildtype control(set as 100%). For RPA, The probe overlaps 5'ss of TRα2 was used to indicate the expression of both TRα1 and TRα2 spliced mRNA. For Realtime PCR assays, The calculated splicing percentage of TRα2 based on the data from primer set A/A1 and B/B1. See Figure 7C and 8B for details.

C. Comparison of the expression of TRα1, TRα2 and downstream spliced products(include cryptic splicing). The result was generated by calculating the corresponding percentage with CT values from primer set A/A1(TRα1) ,B/B1(TRα2) and B/B2(For DS: downstream splicing). CT values of A/A2 is set for the all the transcripts shared with exon8 and exon 9A. The %spliced represents the combined percentage of TRα1, TRα2 and downstream (DS) product with different minigenes. Results of 2 or 3 respective experiments are shown for each constructs.

D. Selected samples were assayed with 2% Agarose gel to test their matched size. The templates and primer sets were indicated in the figure. The lower panel is a high contrast version of these bands. Abbreviation for plasmids used in the Figure: C-control, WT-wildtype, SVLα2-SVLα2, G100-βG100αR100, R100-αG100βR100, GR200-βGR200, GR200+SVLα2-βGR200+SVLα2

products downstream of TRα2 3'ss. Because this primer favors the smaller sized products, it can be used to estimate the relative expression level of TRα3 and cryptic splicing in comparing to selective constructs. Similar to conclusion from RPA and realtime PCR assays, the results suggested that after substitution of ESX10 with Rev-erbβ sequence (GR200), the TRα2 splicing percentage reduced dramatically compared to wildtype control (TRα2- GR200:2.6% compared to WT:20.4%) and TRα1 splicing percentage increased correspondingly as well. (TRα1- GR200:96.3% compared to WT:74.7%) (Figure 9C). Substitution of ESX10 subfragments (G100, R100) moderately reduced TRα2 splicing (TRα2- G100:8% and R100:20.8%). With primer B/B2, splicing products from different splice sites were indicated after PCR amplification (Figure 9D, the panel with primer B/B2). The electrophoresis results indicated that in R100, GR200 and GR200+SVLα2 minigenes, only cryptic spliced products was produced. Because in G100, normal spliced TRα2 is detected but in R100 we observe mainly cryptic splicing products, we suggested that R100 is more important for TRα2 splicing than G100. This statement also agrees with conclusions from previous *in vitro* experiments that R100 fragment showed more enhancing capability than G100 [36].

Although the substitution of ESX10 and its subfragments (R100) induced cryptic splicing, the expression of cryptic splicing (DS- downstream splicing) in R100 and GR200 is much lower compared to normal TRα2 splicing (DS splicing in R100: 0.1% and GR200-0.1%; Compared to the TRα2 splicing in WT: 20.4%). These results suggest that ESX10 and R100 are functional critical for normal TRα2 splicing and the overall splicing level was not reduced dramatically. Since the expression levels of introduced cryptic splicing

products are much lower than TRα2, this conclusion would not be biased by the appearance of cryptic splicing. Interestingly, the expression of downstream splicing after ESX10 substitution was promoted greatly by strong downstream poly(A) signal (DS-GR200:0.1%; GR200+SVLα2:18.2%). In next chapter, the effect of this strong polyadenylation signal from the SV40 virus on splicing of TRα2 will be examined in more details.

C. Discussion:

In this chapter, the functional role of ESX10 and its subfragments (G100 and R100) have been studied through *in vivo* experiments. Constructs with substitution of ESX10 and both subfragments (G100 and R100) have been created and RNase Protection Assays and realtime PCR were used to evaluate the expression levels of normal spliced TRα2 and related cryptic splicing.

Studies of ESX10 subfragments with RPA and realtime assays

In order to study the functional effect of ESX10 subfragments, both RNase Protection Assay and realtime PCR assays were used to investigate TRα alternative splicing. In RNase Protection Assay, the substitution of intact ESX10 with Rev-erbβ sequence resulted a sharp decrease on TRα2 splicing (2.8%) compared to wildtype control WT(20.5%). Substitution of either ESX10 subfragment (G100 and R100) with antisense Rev-erbβ sequence causes reduction of TRα2 splicing (G100:9.0% and R100: 11.5%) (Figure 7B and 7C). Although both G100 and R100 reduced TRα2 splicing, these substitutions were substantially less effective than GR200. The realtime PCR data

showed similar results. Compared to wildtype control WT(16.5%), GR200 resulted a sharp decrease(0.2%) and the respective TRα2 splicing percentage for G100 is 6.0% and R100 is 8.4%(Figure 8B). However, realtime PCR data in Figure 8B measured only TRα2 splicing relative to TRα1 while the RPA results in Figure 7C measure all splicing of the 5'ss in exon 9A. Therefore, both methods suggested that ESX10 sequence is functionally important for TRα2 splicing and its subfragments(G100 and R100) also indicate certain level of enhancing capability and the substitution of either half induced a detectable drop on normal TRα2 splicing. The data also indicated that the G100 substitution caused more decrease on TRα2 splicing than R100 substitution, which suggests that G100 might have more enhancing activity than R100 within the context of the ErbAm minigene. However, this conclusion doesn't agree with previous *in vitro* results from our lab [36]. When R100 and G100 fragments tested *in vitro*, R100 enhanced dsx pre-mRNA to almost same level as GR200 and G100 was much less efficient. Another proof is that in R100, only cryptic splicing product can be detected with electrophoresis and in G100 we observed mainly normal TRα2 product. Therefore, we conclude that although substitution of G100 with Rev-erbβ sequence also causes reduction of TRα2 splicing, R100 is more critical for TRα2 splicing and its substitution will directly inhibit the normal TRα2 splicing. More work is required for studying the expression level of TRα2 in G100 (pErbAm-βG100) to further investigate the effect of G100 on TRα alternative splicing.

Validity of the R100 and G100 minigene constructs

After introducing the substitution of ESX10 subfragments (in G100 and R100 constructs), we detect several nucleotide mutations outside of ESX10 as shown in Appendix A. Compared to rat TRα sequence, one adenine has been mutated into guanine at 160 nt downstream of EcoRI site (the 3' end of ESX10) in G100 construct. In R100 construct, one cytosine has been changed into thymine 210 nt downstream of EcoRI site and one adenine has been changed to guanine in 271 nt downstream of EcoRI site. The appearances of several point mutations within the newly generated constructs are probably due to the error during recombinant PCR process. Although it is possible that the introduced nucleotide changes might induce unpredictable results, we assume that these few scattered mutations have trivial effect on TRα alternative splicing in these minigenes.

Related studies concerning cryptic splicing

As the name describes, cryptic splicing involves splicing at sites not normally used. Cryptic splice sites often resemble to normal splice sites but are not selected until a mutation is introduced that alters usage of authentic splice site. In this Chapter, we identify three cryptic splice sites downstream of TRα2 3' splice site. These include 3' splice site for TRα3, a minor variant which was first describe more than 20 years ago [13] and two sites introduced by Rev-erbβ substitution. The introduced cryptic splicing site in G100 is very close to the TRα3 splicing site in original TRα sequence and therefore a product have similar size to TRα3 was observed. In contrast, the R100 (pErbAm-βR100) shares a similar cryptic splicing site in substituted R100 with GR200 (pErbAm-β200). As

can be seen in Table I, these are all slightly poorer matches to the mammalian consensus sequence. Interestingly, the two sites in the Rev-erbβ sequence both resemble the normal TRα2 3'ss in that they have a AAG/G splice site sequence (Table I).

In our system, although the introduced cryptic splicing is close to the authentic TRα2 splicing site, the sequence for normal TRα2 3' splicing site has not been mutated. This raises a challenging question related to the function of ESX10: does cryptic splicing occur because of the missing of correct ESX10 sequence or does introduction of cryptic splice sites compete with the normal splicing of TRα2? For cryptic splicing within in GR200 and R100, the introduced cryptic splice site might not so strong as to compete efficiently with the authentic TRα2 3'ss. Rather the absence of the enhancer activity of ESX10 may change the splicing pathway for TRα2 5'ss. On the other hand, the cryptic splice site within the G100 substitution is very close to the original TRα3 splicing location. It is possible that (undefined) elements that direct splicing of TRα3 results in selection another usable site near the original site. Although the importance of cryptic splicing is trivial compared to the ESE studies, the selection between different sites (strong or weak) might provide reasonable explanation for the interaction between cis-acting elements and normal splicing factors.

<u>The endogenous TRα2 in transfected HEK293 cells affect the realtime PCR results</u>

One problem we encountered in the realtime PCR process is in transfected HEK cell or cell transfected with vector as negative controls, we found a positive signal for TRα2 splicing. Although the expression level is much lower compared to the experimental ones,

it would affect the measurement of minor splicing from analysis of realtime data. One hypothesis is that endogenous TRα2 spliced in human cells would also be detected and included in the C_T values and low levels of TRα2 might represent cross contamination. In order to solve this problem, we isolated the TRα2 PCR product obtained from untransfected cells and compare the sequencing result from this sample with human and rat sequence. Since the exon10 of TRα2 is highly conserved between these two species, we used the nucleotides difference among exon 8 and exon 9 as major criteria for our judgments and overall five nucleotides were accepted during the process. The sequence results demonstrate the existence of endogenous TRα2 from human cell (Appendix-B). Therefore, the C_T values from realtime PCR correctly represent the expression level of TRα2 and not adventitious contamination or some other artificial signal.

Chapter IV. Effect of strong polyadenylation on TRα2 alternative splicing

A. Introduction

Polyadenylation is an important post-transcriptional process for eukaryotic mRNAs. This chapter focuses on the contribution of polyadenylation to the regulation of TRα2 splicing. In our system, the sequence of simian virus 40 late (SVL40) mRNA is used as a strong polyadenylation signal, and the effect of this strong downstream poly(A) on TRα2 and TRα3 expression was studied. We also investigated TRα2 splicing with a strong downstream poly(A) signal when the ESX10 was substituted. Finally, the effect of strong poly(A) signal on TRα2 splicing was also studied when authentic TRα2 5'ss was disrupted.

B. Analysis of TRα3 splicing demonstrates that downstream SVL poly(A) signal enhances the efficiency and fidelity of normal TRα2 splicing.

When using pCMVerbAα minigene as a wildtype control, TRα3, a different isoform of TRα2 with a shorter size was detected by PCR (Figure 10B). Sequencing of the small PCR product seen here and Table I also indicated that this product represents TRα3 3'spliced site, which is 117 nt downstream of the TRα2 3'ss (Figure 10A, Figure 4 and Table I). Next, a known strong poly(A) site, late SV40 was inserted in place of downstream TRα2 poly(A) site to test the effect of strong pA on TRα3 splicing. Surprisingly, in the construct that has strong TRα2 poly(A) signal, the expression of TRα3 was reduced dramatically in comparison with the TRα2 mRNA (Fig 10B). Realtime PCR results with primer specific for TRα2 splicing also showed that the strong

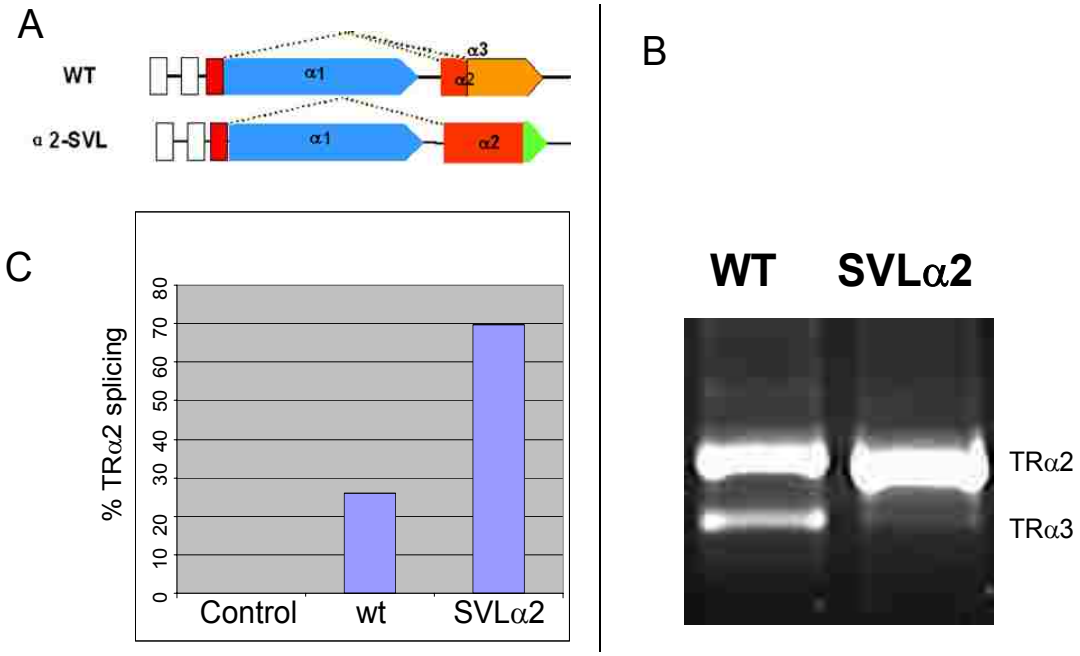


Fig 10. Introduction of strong poly(A) site SVLα2 increase TRα2 splicing fidelity and eliminate TRα3 expression

A. Schematics structure of TRα2 and TRα3
B. Regular PCR results confirm an elimination of TRα3 expression and increase on TRα2 expression. A/A1 and B/B1 were used.
C. Real-time data indicate that TRα2 splicing increases with strong downstream polyadenylation. The results are from calculation of TRα2 percentage from a combination of TRα1 and TRα2. (See Chapter V for more details)

poly(A) signal enhances the expression level of regular TRα2 splicing compared to TRα1 splicing (Figure 10C). Thus it appears that inserting a strong poly (A) signal downstream of TRα2 increases the splicing fidelity of correct TRα2. Expression level of TRα3 isoform would be inhibited and TRα2 splicing is enhanced.

C. Effect of the Strong Poly(A) Signal on Cryptic Splicing

1. The effect of downstream strong poly(A) signal on TRα2 splicing when normal TRα2 5'ss is disrupted

In an effort to determine wether the strong poly (A) for TRα2 can compete directly with that for TRα1 in the absence of TRα2 splicing, transfection were carried with a minigene construct that lack the TRα2 5'splice site (pErbAm Δ5'ss) paired with downstream strong polyadenylation site of TRα2 (pErbAm Δ5'ss+SVLα2). It was expected that the major products would be two RNAs differing only in their choice of poly(A) sites. In pErbAm Δ5'ss, six nucleotides were changed compared to the TRα2 5'ss consensus sequence (Figure 11A) and previous results from our lab indicated that this mutation of TRα2 5'ss (Δ5'ss) completely blocks TRα2 splicing [44]. Conventional PCR results suggested that disturbance of TRα2 5'ss also inhibited the expression level of TRα3(Figure 11B). TRα2 splicing is not enhanced by the strong TRα2 poly (A) signal when authentic 5'ss was disturbed (pErbAmΔ5'ss+SVLα2). However, cryptic splicing products were detected in pErbAmΔ5'ss and promoted by the SVL poly (A) signal (pErbAmΔ5'ss+SVLα2) (Figure 11B). After sequencing the respective bands, two cryptic splice sites have been identified located 53 nt and 142 nt downstream of normal TRα2 5'ss in exon 9 (Fig 11A). The more distal site (CSS2) did not yield a clear sequence suggesting that the large

A

CAG/ACCGCTCTGGCCTGCTGTGTGTGGACAAGATCGAGAAGAGTCAGGAGGCCTACCTGCTGGCGTT
3'ss

△5'ss: CC/CAC
TGAGCACTACGTCAACCACCGCAAACACAACATTCCGCACTTCTGGCCCAAGCTGCTGATGAAG/GTG

TRα2 5'ss(exon 9A)

AC GA
ACTGACCTCCGCATGATCGGGGCCTGCCACGCCAGCCGCTTCCTCCACAT/GAAAGTCGAGTGCCCCA

CSS1

CCGAACTCTTCCCCCCACTCTTCCTGGAGGTCTTTGAGGATCAGGAAGTCTAAAGCCTCAGGCGGCCA

TRα1 stop

GAGG/GTGTGC
CSS2

B

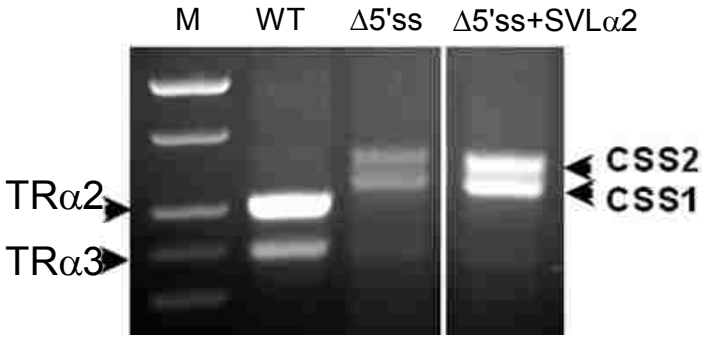


Figure 11 Strong poly(A) site activates weak splice site near TRα2 5'ss

A. Sequence analysis of exon9 in TRα. Different splicing sites are marked: ss stands for splicing site, CSS1 and CSS2 represents cryptic splicing sites 1 and 2. Sequence substitutions in 5'ss of pErbAm △5'ss are shown above.
B. Conventional PCR amplification after reverse transcription are used to check TRα2 splicing. With constructs having mutated 5'ss for on exon 9, a decrease of TRα2 has been detected. A strong poly (A) site will enhance the use of available weak splicing site and increase the expression of cryptic spliced products. The primer set used to check cryptic splicing is 936F/905R.

PCR products may actually represent a mix of two (or more) products spliced a few nucleotides apart. It appears that a strong downstream poly (A) signal enhances the usage of weak splice site near the mutated TRα2 5'ss and increase cryptic splicing (Figure 11B).

2. <u>The effect of downstream strong polyA signal on TRα2 splicing in the absence of ESX10</u>

As we discussed in the Chapter III, normal TRα2 splicing was reduced greatly after ESX10 sequence substitution. Previous studies also showed that strong polyadenylation site could enhance TRα2 splicing activity (Figure 12A, Lanes WT and SVLα2). Therefore, a series of tests have been made to study whether a downstream strong poly (A) signal will also induce positive effects on TRα2 splicing in the absence of ESX10. By comparing the expression level of TRα1 and TRα2, we hope to investigate the interactions between ESX10 and polyadenylation mechanism.

In order to study the effect of downstream strong poly (A) signal on TRα2 splicing in the absence of ESX10, a strong poly (A) has been introduced into the construct that has ESX10 substituted with corresponding Rev-erbβ sequence. Conventional PCR and realtime assays were used to analyze the expression level of TRα2 and other spliced product. The results from regular PCR indicated that normal TRα2 splicing was lost by substitution of Rev-erbβ sequence in ESX10 and the additional strong SVL40 poly (A) signal didn't increase normal TRα2 splicing (Figure 12A). The usage of introduced weak splicing sites near TRα2 3'ss were, however, enhanced after introducing downstream

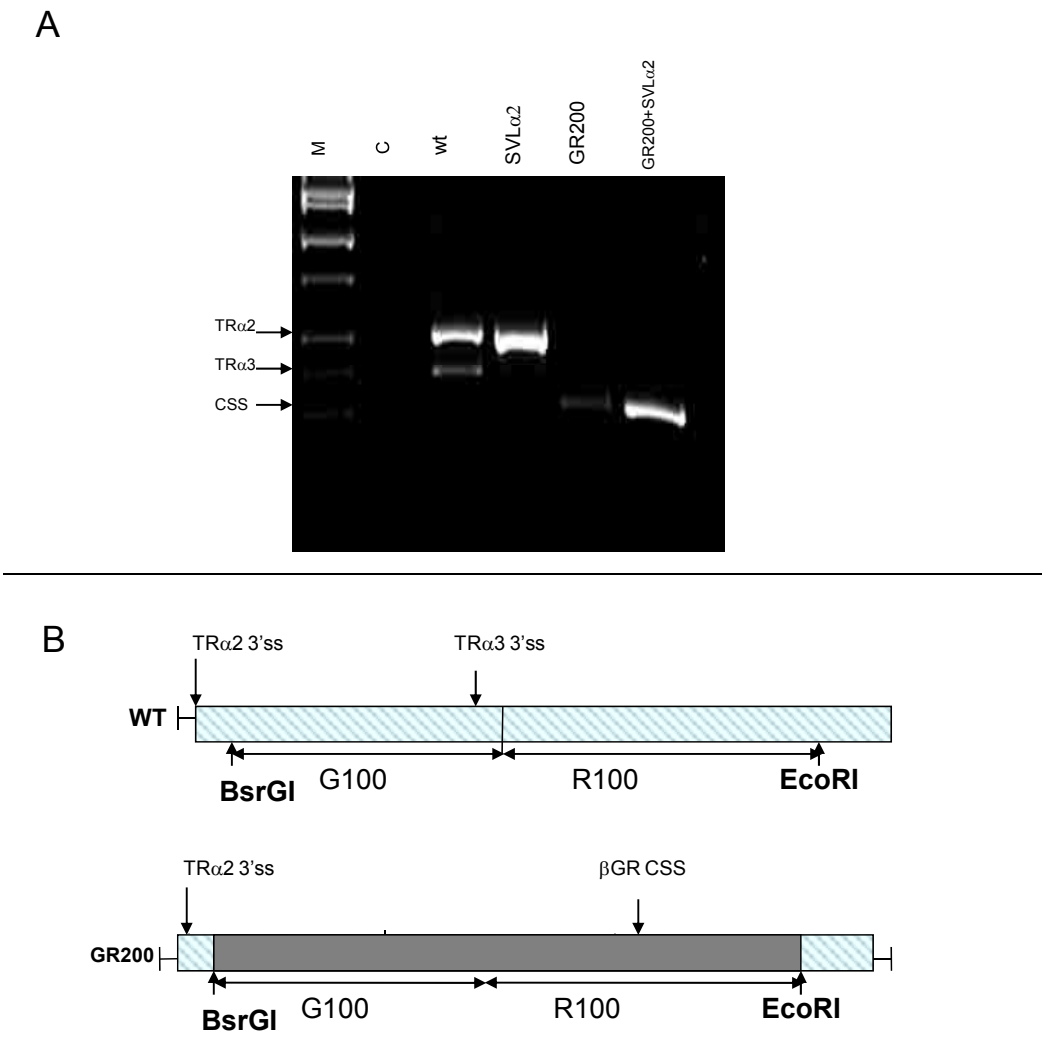


Figure 12 Downstream polyadenylation activates cryptic splicing at TRα2 3'ss

A. Regular PCR bands indicate thatESX10 are critical for TRα2 correct splicing and strong poly(A) signals will enhance general splicing level and increase correct splicing with the existence of ESX10 enhancer.

B. The schematic structure of different splicing sites of wt and GR200 from sequence analysis and spliced product size.

poly (A) site (Figure 12 B). The identity of the enhanced cryptic splicing products was confirmed by sequencing.

D. Discussion:

In this chapter, we further studied the relationship between downstream polyadenylation site strength and TRα2 alternative splicing. Previous results indicated that the downstream poly(A) will promotes the nearby splicing and my results indicate that the strong downstream poly (A) signal will not only enhance TRα2 splicing but increase the fidelity of TRα2 splicing by reducing TRα3 splicing. One possibility is that ESX10 mediates the definition of exon 10, which may also explain why addition of the SVL site eliminates TRα3 splicing. However, insertion of the strong poly (A) signal also activates cryptic splicing at two sites near the 5'ss upstream in exon 9. In the absence of normal TRα2 5'ss, these cryptic spliced products are promoted with TRα2 downstream polyadenylation signal. With the antisense Rev-erbβ substitution for ESX10, the usage of introduced weakly cryptic splicing sites is enhanced by the strong SVLα2 replacement at TRα2 poly (A) site. Although the cryptic sites associated with ESX10 are only technical problems, it confuses our analysis on the ability of the Rev-erbβ substitutions to block normal TRα2 splicing.

Previous experiments showed that disruption of TRα2 5'ss by the mutation (Δ5'ss) would completely inhibit the TRα2 slicing. However, no cryptic splicing or TRα3 products were observed in a previous study using northern blot analysis, presumably because that technique is less sensitive than PCR [44]. In my experiments, the cryptic splicing was

detected during the process of studying the effect of TRα2 downstream poly(A) on TRα2 splicing. With the primer A/B2, additional PCR products that are larger than those for TRα2 and TRα3 were detected. After sequencing, we concluded that elimination of the 5'ss for TRα2 induces cryptic splicing at site 53 and 142 nucleotides downstream of the mutant site (Table II). And these cryptic splicing were weak in the Δ5'ss without SVLα2 poly (A) site but their expression levels were greatly enhanced by TRα2 downstream poly (A) (Figure 11C). The sequence for cryptic splicing site 1(CSS1) was proved consistent with reproducible experiments. However, the splicing site for CSS2 was not located consistently and several splicing sites were identified by separate sequencing data (CSS2 in Table II represent one confirmed CSS2 splicing sites). Therefore, we assume that there might be multiple cryptic splicing sites available near the identified CSS2 site and different products with similar size would be produced after disrupting correct TRα2 5'ss.

In this chapter, the presence of the cryptic sites associated with the strong SVL polyA site were identified and investigated by various approaches. Although previous studies indicate that the expression level of these cryptic splicing products is much lower than normal TRα splicing(Figure 9C), the existence of these cryptic splicing complicate analysis of possible interactions between the poly(A) site and the ESX10 splicing enhancer. Further work to eliminate the expression of cryptic splicing in our system would be helpful for characterization of the interaction between ESX10 and polyA and their effect on TRα alternative splicing.

Table II. Sequence of TRα2 5'ss and other cryptic splicing sites

	position	sequence
Consensus	N/A	AAG/GURAGU
TRα2 5'ss	+1	AAG/GUGACU
TRα2 5'ss CSS1	+53	CAU/GAAAGU
TRα2 5'ss CSS2	+142	AGG/GUGUGC

In this table, R stands for purine nucleotides. 5'ss and CSS respectively represent 5' splicing site and cryptic splicing sites.

V. Discussion and Conclusion

Advantage and limitation of mutation studies

The characterization of ESX10 and its subfragments was investigated by mutation studies. These studies are normally carried out to further narrow down the key regulatory region in a defined large fragment. If the authentic sequence contributes significantly to the enhancing effect of ESX10, mutation of the target sequence would cause an increase or decrease in the amount of splicing. However, if there is no effect after substitution, the region might not be as important as hypothesized. In our system, TRα minigene provides an good model to further study the regulatory elements, like enhancers within exon or intron [32, 34, 36, 37].

The major limitation of this method is that mutations *in vivo* sometimes will activate cryptic splicing sites either within mutated region or near mutated authentic splicing site. In one study from Krainer and colleagues, the sequence difference between authentic 5' splice sites and cryptic splice sites has been investigated [48]. Different statistical scoring methods were used to compare the strength of authentic 5' splice sites and cryptic splicing sites. The authentic 5' splice sites ranked highest among all the accepted methods and the mutant sites ranked lowest. The cryptic splice sites received intermediate score compared to the authentic 5'ss and mutant 5'ss. Interestingly, the activated cryptic splicing sites were not necessarily the best scoring alternative sites near the authentic 5' splicing sites, which suggests that other factors could also contribute to cryptic splicing sites selection. For example, relative usage of cryptic splicing sites near normal 5'ss would be affected by increasing SR protein or hnRNP A/B in mutant

substrates of *Hbb*(human β-globin gene) [49, 50]. In our system, the cryptic splicing activated after disrupted TRα2 5'ss could be due to the loss of authentic splicing site. And the cryptic splicing happened within substituted ESX10 or its subfrgaments (G100 and R100) might be due to the dysfunction of ESX10 and existence of other factors involved in alternative splicing. Further efforts on identifying the trans-acting factors would be necessary to characterize the mechanisms involved with ESX10 and its subfragments.

The advantages and disadvantages of *in vivo* and *in vitro* methods

In this thesis, I applied both *in vivo* and *in vitro* methods to investigate the identification and characterization of cis-acting elements for TRα2 splicing. The major advantages of Dsx *in vitro* splicing system are its sensitivity and quick application. Since this splicing system lacks its original enhancer and relies totally on the introduced enhancing segments, even an enhancer with low level could be investigated and slight change on the enhancing capability could also be easily detected. In R100, three ESE candidates have been predicted by Rescue-ESE program and their activities were tested through *in vitro* experiments (Figure 5). After testing, the R1 ESE motif seems to be the most efficient one among these three motifs and we can conclude that one important factor that could contribute significantly for the ESX10 activity and following TRα2 splicing. However, in the following *in vitro* competition experiments, the fragment with substituted R1, R2 and R3 did not compete against the pre-mRNA splicing as efficiently as authentic GR170 fragments (Figure 6C). This suggests that the mutated fragments don't bind some factors essential for TRα2 splicing. Combining the *in vitro* and *in vivo* results, we cannot

determine whether the segments overlapping R1 R2 R3 are necessarily the only ones required for TRα2 splicing. Further work would include narrowing the competitor fragments(from current GR170 to R70) and detecting trans acting factors.

Although *in vitro* experiments are very sensitive for detecting possible enhancer elements, there are also disadvantages: in the *in vitro* system, the effect of certain factors inside the cell might be neglected and their absence or perturbation may affect the enhancing capability in real situation and biased the accuracy of our conclusion. Therefore, *in vitro* splicing test is useful approach for analyzing the activity of enhancers and *in vivo* studies would be helpful to confirm the results convincingly. Thus we further characterize the functional importance of ESX10 subfragments with a series of *in vivo* studies after creating constructs with substitution of either G100 or R100. The results from RNase Protection Assay and realtime PCR suggested that both ESX10 subfragments had enhancing capability on TRα2 splicing and substitution of either G100 or R100 would caused the decrease on TRα2 splicing (Figure 7 and 8). On the other hand, the substitution of G100 reduced TRα2 splicing more effectively that the R100 substitution, which suggests that factors associated with G100 may also contribute significantly to TRα alternative splicing. However, previous *in vitro* experiments confirmed that R100 had more enhancing capability than G100, which against the conclusion we draw through RPA and realtime assays.

The contradictions between the enhancing capability of ESX10 subfragments in *in vitro* and *in vivo* experiments may be due to several reasons. First, when we test the activity of

G100 and R100 through *in vitro* experiments, the enhancing capability cis-acting elements and trans-acting factors on the boundary of the subfragments (upstream sequence of G100, boundary between G100 and R100, downstream sequence of R100) would be affected and their original function might be significant. Second, although the dsx *in vitro* experiments are sensitive approach for detecting enhancers, minor difference between *in vitro* and *in vivo* experiment (ie. the *in vitro* incubation environment and *in vivo* cell growing) could induce unexpected effect on the final results. Therefore, further efforts on identifying what proteins bind to ESX10 is necessary to better understand the mechanism and solve the problem. Combining the results from both *in vivo* and *in vitro*, we further characterize the functional role of ESX10 subfragments and increased our understanding on the regulatory mechanism of ESX10.

Studies related to Polyadenylation and cis-acting elements

Generally, cis-acting elements normally work as the docking sites or targets for trans-acting factors and affect the activity of the nearby splice sites [51, 52, 53]. Polyadenylation has also been confirmed to affect splicing through protein-protein interactions among polyadenylation factors and splicing factors, where downstream polyadenylation positively affect the activity of proximal splice site [54]. Splicing and polyadenylation factors also interact with RNA polymerase II and other interactions before transcription and processing are also likely to occur. Therefore, the next goal was to investigate the possible interactions between cis-acting elements and downstream poly (A) signals.

Recent studies also pointed to important interactions between chromatin structure and alternative splicing. For example, studies of the fibroblast growth factor receptor2 (FGFR2) gene illustrate a possible relationship between histone modification and alternative splicing may provide insight into our results [55]. In this study, distinctive histone modifications were shown to correlate with the effects of cis-acting elements on the splicing outcome in a series of human genes. The recruitment of the splicing regulatory protein PTB that binds to RNA would be affected by a chromatin-binding protein MRG15. Thus alternative spicing events may be regulated through this adapter system. Given these results, it is quite possible that polyadenylation factors also directly or indirectly interact with similar components of chromatin. If such interactions can be confirmed, the bridge between cis-trans interaction mechanism and polyadenylation mechanism could be established.

Conclusions and Summary

In this thesis, we explored the role of ESX10 and the downstream poly (A) on promoting or facilitating TRα2 splicing. Three possible ESE motifs within ESX10 were identified and their enhancing effects on TRα2 splicing were investigated *in vitro*. Following *in vivo* approaches were applied to further characterize the functional importance of ESX10 subfragments. In the end, we studied the effect of strong downstream poly (A) on TRα2 splicing and other related splicing events (TRα3 and cryptic splicing). The results from these experiments suggest that the existence of cross-exon interactions involve three elements: the 3'ss, the ESEs (ESX10) and the pA site. This network of interactions may be targeted or affected by antisense RNA produced through transcription of the opposite

strand (Rev-erbα) indirectly, by altering the modification of the chromatin. The roles of TRα1 and Rev-erbα as two regulatory proteins important for developmental regulation, differentiation and metabolic control suggest that expression of this locus may be regulated in many different ways. TRα2 might function as an additional layer of regulation that modulates expression of these genes. The physiological role of TRα2 is also itself remarkably elusive. More than 20 years after its discovery the function of this protein is unknown despite the fact it is highly conserved and its expression is differentially regulated in different cells and tissues of eutherian mammals. This study may help to expand our current knowledge on regulatory mechanism of TRα alternative splicing and the possible interaction between TRα2 and other thyroid hormone receptor proteins.

VI. Materials and Methods

Plasmids

Constructs with ESX10 Subfragments Substitution

To further investigate the functional effect of subfragments of ESX10, substitution of either halves of ESX10 was introduced into minigene constructs. The parent vector used was 633(pErbAm delta 2+SVLα2), which has a 200nt truncation downstream of EcoRI site on exon10. The mutated G100 (βG100αR100) and R100 (αG100βR100) insertion fragments were created by recombinant PCR. For G100, primer 979F and 1493R were used to create upstream βG100 fragment and 982R and 1492F were used for downstream αR100 fragment. The recombinant βG100αR100 fragment was generated with 979F and 982R. For R100 construct, Primer 979F and 1491R were used to create upstream αG100 fragment and 982R and 1490F was used to generate downstream βR100 fragment. Similarly, the recombinant αG100βR100 fragment was generated with 979F and 982R as well. Both parent vector and the insertion fragments were digested with BsrGI and SbfI restriction enzymes and ligated by T4 DNA Ligase after digestion. Since the insertion introduced a 200 nt longer fragment into the parent vector, the mutations were confirmed by Bsu36I restriction enzyme digestion and sequencing results. During the sequencing process of the new constructs, several point mutations have been identified downstream of ESX10 (Appendix C). Although we didn't detect any obvious cis-acting elements near the region where these mutated nucleotides were identified, these changes might introduce unpredicted effect on the potential regulatory mechanisms near these regions.

Other used plasmid can be found in Appendix-D.

Cell lines: HEK293 cells (also referred as Human Embryonic Kidney 293), is derived from tissue culture of human embryonic kidney cells and stored in our lab.

Transfection and RNA isolation

The calcium phosphate mediated transfection was set up according to the standard procedure of Sambrook et al (1989). HEK293 cells were plated as concentration of 2×10^6 per P100 plate 24 hours before transfection and 48 hours before RNA extraction. After one day, 5 ug of plasmid DNA calcium phosphate were transfected into each plate. The cells were rinsed 24 hours after transfection and fresh media was introduced. The RNA extraction was performed with RNA extraction buffer and 2x PK buffer. DNaseI and PK were used in the process to purify and concentrate the target RNAs. RNA extraction from part of the transfection undergoes Trizol prep extraction (TRIzol Isolation of RNA, DNA and Protein, Invitrogen).

RNase Protection Assay

To obtain the TRα1/TRα2 mRNA alternative splicing ratio, specific probes were designed according to the corresponding sequence. The DNA used to produce the radio-labeled probes was generated by PCR with designed primers. Different probes were generated according to different spliced product and cryptic splicing. Various concentrations of RNA and probes were hybridized overnight. RNase Cocktail containing RNases A and T1 and following Proteinase K treatment were used to stop the reaction. The samples were extracted, ethanol precipitation, resuspended in FSB and load on a 5.5% polyacrylamide-urea gel.

In vitro transcription and splicing

Most RNA substrates were generated with PCR templates. Different types of polymerase (T7 or SP6) were used according to promoter specificity. After Mix and incubate at 37 ºC for 60 minutes. DNase I was used to eliminate the residual DNA templates. After overnight EtOH precipitation, RNA pellets will be resuspended in 20 ul FSB and load on to polyacrylamide-urea gel to further isolate. Otherwise, equal volume of urea gel stop buffer can be added after *in vitro* transcription reaction and the mixture will be loaded onto a polyacrylamide gel. After treatment with elution buffer for 4 hours, phenol extraction and following precipitation are used to purify the RNA. For *in vitro* splicing reaction, nuclear extract was incubated with pre-mRNAs, and a certain portion of reaction was taken out after a fixed time period and incubates with PK for 25min. After precipitation, labeled RNA was resuspended in FSB and load to 5.5% gel.

Quantitation of assays

The dried gel after RPA and *in vitro* splicing were analyzed by scanning with Storm phosphor-imager to collect counts of each detectable band. Because the amount of unlabeled UTP and labeled hot UTP were preset before the reaction, the different number Us in different product are used to measure the actual percentage of spliced mRNA against unspliced product.

Quantitation of mRNA expression level from realtime PCR

After converting mRNA into complementary DNA with reverse transcriptase, we use realtime PCR to investigate the expression level of mRNAs. Since DNA amount doubles

after every cycle of PCR until it reaches the corresponding thresholds, the Ct values can be used to calculate the percentage of each product(TRα1, TRα2 or cryptic splicing product). . And the difference between TRα1 and TRα2 Ct values was used to measure TRα1/ TRα2 ratio. The lower Ct value is, the higher the initial mRNA presents in the reaction. And the specific calculation shows below:

In Figure 8B, we use primer set A/A1 for TRα1 and B/B1 for TRα2 detection

$$\%TR\alpha2 \text{ splicing} = 100*TR\alpha2/(TR\alpha1+TR\alpha2) = R*100/(R+1)$$

$$[R = TR\alpha1/TR\alpha2 = 2^{Ct(B/B1)-Ct(A/A1)}]$$

In Figure 9B, we use primer set A/A1 for TRα1, B/B1 for TRα2, B/B2 for CSS and A/A2 to detect the expression of all the transcripts with shared exon8 and exon 9A.

$$\%(TR\alpha1+TR\alpha2+CSS) = 100*(TR\alpha1+TR\alpha2+CSS)/(\text{all transcripts})$$

$$[TR\alpha1/(\text{all transcripts}) = 2^{Ct(A/A2)-Ct(A/A1)}]$$

$$[TR\alpha2/(\text{all transcripts}) = 2^{Ct(A/A2)-Ct(B/B1)}]$$

$$[CSS/(\text{all transcripts}) = 2^{Ct(A/A2)-Ct(B/B2)}]$$

In Figure 9C, the expression level of TRα1 and TRα2 were calculated.

$$\%TR\alpha1 = 100*TR\alpha1/(TR\alpha1+TR\alpha2+CSS)$$

$$\%TR\alpha2 = 100*TR\alpha2/(TR\alpha1+TR\alpha2+CSS)$$

VII. BIBLIOGRAPHIES

1. Black DL. Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem 2003; 72:291-336.

2. Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet 2002; 3:285-98.

3. Alt, FW. Bothwell AL, Baltimore D *et al.* Synthesis of secreted and membrane-bound immunoglobulin μ heavy chains is directed by mRNAs that differ at their 3' ends. Cell 1980; 20: 293–301.

4. Early, P. Rogers J. Hood L *et al.* Two mRNAs can be produced from a single immunoglobulin μ gene by alternative RNA processing pathways. Cell 1980; 20: 313–319.

5. Pan, Q., Shai, O., Lee, L. J., Frey, B. J. & Blencowe, B. J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nature Genet 2008; 40: 1413–1415

6. Wang, E. T. Sandberg R. Burge CB. *et al*. Alternative isoform regulation in human tissue transcriptomes. Nature 2008; 456: 470–476.

7. Nilsen TW and Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. Nature 2010; 463: 457-463

8. Smith,CW and Valcarcel,J. Alternative pre-mRNA splicing: the logic of combinatorial control. Trends Biochem Sci 2000; 25: 381–388.

9. Lazar MA, Hodin RA, Darling DS, Chin WW. A novel member of the thyroid/steroid hormone receptor family is encoded by the opposite strand of the rat c-erbA alpha transcriptional unit. Mol Cell Biol 1989; 9:1128-36.

10. Lazar MA, Hodin RA, Cardona G, Chin WW. Gene expression from the c-erbA alpha/Rev-ErbA alpha genomic locus. Potential regulation of alternative splicing by opposite strand transcription. J Biol Chem 1990; 265:12859-63.

11. Strait KA, Schwartz HL, Perez-Castillo A, Oppenheimer JH. Relationship of c-erbA mRNA content to tissue triiodothyronine nuclear binding capacity and function in developing and adult rats. J Biol Chem 1990; 265:10514-21.

12. Munroe SH, Lazar MA. Inhibition of c-erbA mRNA splicing by a naturally occurring antisense RNA. J Biol Chem 1991; 266:22083-6.

13. Mitsuhashi T, Tennyson GE, Nikodem VM. Alternative splicing generates messages encoding rat c-erbA proteins that do not bind thyroid hormone. Proc Natl Acad Sci U S A. 1988; 85:5804-8.

14. Lazar MA. Thyroid hormone receptors: multiple forms, multiple possibilities. Endocr Rev. 1993; 14:184-93.

15. Lazar MA, Hodin RA, Darling DS, Chin WW. A novel member of the thyroid/steroid hormone receptor family is encoded by the opposite strand of the rat c-erbA alpha transcriptional unit. Mol Cell Biol 1989; 9:1128-36.

16. Miyajima N, Horiuchi R, Shibuya Y, Fukushige S, Matsubara K, Toyoshima K, et al. Two erbA homologs encoding proteins with different T3 binding capacities are transcribed from opposite DNA strands of the same genetic locus. Cell 1989; 57:31-9.

17. Yin L, Wu N, Curtin JC, Qatanani M, Szwergold NR, Reid RA, et al. Rev-erbalpha, a heme sensor that coordinates metabolic and circadian pathways. Science 2007; 318:1786-9.

18. Wang J, Lazar MA. Bifunctional role of Rev-erbalpha in adipocyte differentiation. Mol Cell Biol 2008; 28:2213-20.

19. Drabkin, H. A, Kao, F.-T, Weinberger, C, Evans, R. Human c-erb-A is located at chromosome 3p21.33-p22. Am J Hum Genet 1987; 41: A25.

20. Lazar, M. A.; Chin, W. W. Nuclear thyroid hormone receptors. J Clin Invest 1990; 86: 1777-1782.

21. Iskaros, J.; Pickard, M.; Evans, I.; Sinha, A.; Hardiman, P.; Ekins, R. Thyroid hormone receptor gene expression in first trimester human fetal brain. J Clin Endocr Metab 2000; 85: 2620-2623.

22. Yang, Z.; Privalsky, M. L. Isoform-specific transcriptional regulation by thyroid hormone receptors: hormone-independent activation operates through a steroid receptor mode of coactivator interaction. Molec Endocr 2001 ; 15: 1170-1185.

23. Modrek B, Lee C. A genomic view of alternative splicing. Nat Genet 2002;30:13-9.

24. Dreyfuss, G., Kim, V. N. & Kataoka, N. Messenger-RNA-binding proteins and the messages they carry. Nature Rev Mol Cell Biol 2002; 3: 195–205.

25. Chen M, Manley JL. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nat Rev Mol cell 2009; 10:741-754

26. Matlin AJ, Clark F, Smith CW. Understanding alternative splicing: towards a cellular code. Nat Rev Mol Cell Biol 2005; 6:386-98.

27. Gabut, M., Chaudhry, S. & Blencowe, B. J. The splicing regulatory machinery. Cell 2008; 133: 192.

28. Millevoi S, Vagner S. Molecular mechanisms of eukaryotic pre-mRNA 3'end processing regulation. Nucleic Acids Research 2009; 38:1–18

29. Yeo, G. W. *et al.* An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. Nature Struct Mol Biol 2009. 16, 130–137.

30. Mauger, D. M., Lin, C. and Garcia-Blanco, M. A. hnRNP H and hnRNP F complex with Fox2 to silence fibroblast growth factor receptor 2 exon IIIc. Mol Cell 2008; 28:5403-19

31. Tagami T, Kopp P, Johnson W, Arseven OK, Jameson JL. The thyroid hormone receptor variant alpha2 is a weak antagonist because it is deficient in interactions with nuclear receptor corepressors. Endocrinology 1998; 139:2535-44.

32. Hastings ML, Wilson CM, Munroe SH. A purine-rich intronic element enhances alternative splicing of thyroid hormone receptor mRNA. RNA 2001; 7:859-74.

33. Hastings ML. *et al* Expression of the thyroid hormone receptor gene, erbAalpha, in B lymphocytes: alternative mRNA processing is independent of differentiation but correlates with antisense RNA levels. Nucleic Acids Res 1997; 25:4296-300.

34. Hastings ML. Regulation of erbAa Alternative pre-mRNA Processing. Ph.D Thesis 1998, Marquette University, Milwaukee, WI.

35. Xing Y, Lee C. Alternative splicing and RNA selection pressure--evolutionary consequences for eukaryotic genomes. Nat Rev Genet 2006; 7:499-509.

36. Salato VK, Rediske NW, Zhang C, Hastings ML and Munroe SH. Exonic splicing enhancer within a bidirectional coding sequence regulates alternative splicing of an antisense mRNA. RNA biology 2010 ; 7: 179-190.

37. Salato VK. Sequence elements that regulate splicing of Thyroid hormone receptor mRNA. Master Thesis 2003, Marquette University, Milwaukee, WI

38. Tanaka K, Watakabe A, Shimura Y. Polypurine sequences within a downstream exon function as a splicing enhancer. Mol Cell Biol 1994; 14:1347-54.

39. Edmonds M. A history of poly A sequences: from formation to factors to function. Prog Nucleic Acid Res Mol Biol 2002; 71:285-389.

40. Lutz CS. Alternative polyadenylation: a twist on 3' end formation. ACS Chemical Biol 2008; 3:609-17.

41. Tian B, Hu J, Zhang H, Lutz CS. A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucl Acids Res 2005; 33:201-12.

42. Kiyosawa H, Mise N, Iwase S, Hayashizaki Y, Abe K. Disclosing hidden transcripts: mouse natural sense-antisense transcripts tend to be poly(A) negative and nuclear localized. Genome Res 2005; 15:463-74.

43. Fairbrother WG, Yeo GW, Yeh R, Goldstein P, Mawson M, Sharp PA, et al. RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons. Nucleic Acids Res 2004; 32:W187-90.

44. Hastings ML, Ingle HA, Lazar MA, Munroe SH. Post-transcriptional regulation of thyroid hormone receptor expression by cis-acting sequences and a naturally occurring antisense RNA. J Biol Chem 2000; 275:11507-13.

45. Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: A web resource to identify exonic splicing enhancers. Nucleic Acids Res 2003; 31:3568-71.

46. Stamm S, et al. ASD: a bioinformatics resource on alternative splicing. Nucleic Acids Res 2006 ; 34: 46–55.

47. Wang, Z., Rolish, M. E., Yeo, G., Tung, V., Mawson, M. and Burge, C. B. Systematic identification and analysis of exonic splicing silencers. Cell 2004; 119: 831-845.

48. Roca X, Sachidanandam R, Krainer AR. Intrinsic differences between authentic and cryptic 5' splice sites. Nucleic Acids Res 2003; 31:6321-33.

49. Smith PJ, Zhang C, Wang J, Chew SL, Zhang MQ, Krainer AR. An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. Hum Mol Genet 2006; 15:2490-508.

50. Mayeda A, Screaton GR, Chandler SD, Fu XD, Krainer AR. Substrate specificities of SR proteins in constitutive splicing are determined by their RNA recognition motifs and composite pre-mRNA exonic elements. Mol Cell Biol 1999; 19:1853-63.

51. Buratowski,S. Connections between mRNA 3' end processing and transcription termination. Curr Opin Cell Biol 2005; 17: 257–261.

52. Krainer,A.R., Conway,G.C. and Kozak,D. The essential premRNA splicing factor SF2 in splice site selection by activating proximal sites. Cell 1990; 62: 35-42.

53. Mayeda,A. and Krainer,A.R. Regulation of alternative premRNA splicing by hnRNP A1 and splicing factor SF2. Cell 1992; 68: 365-375

54. Minvielle-Sebastia L, Keller W. mRNA polyadenylation and its coupling to other RNA processing reactions and to transcription. Curr Opin Cell Biol 1999; 11:352-7.

55. Reini F. Luco, *et al.* Regulation of Alternative Splicing by Histone Modifications. Science 2010 ; 327: 996

56. Sambrook and Russel. Molecular Cloning-A laboratory manual.Cold Spring Harbor Laboratory Press; 3rd edition (January 15, 2001)

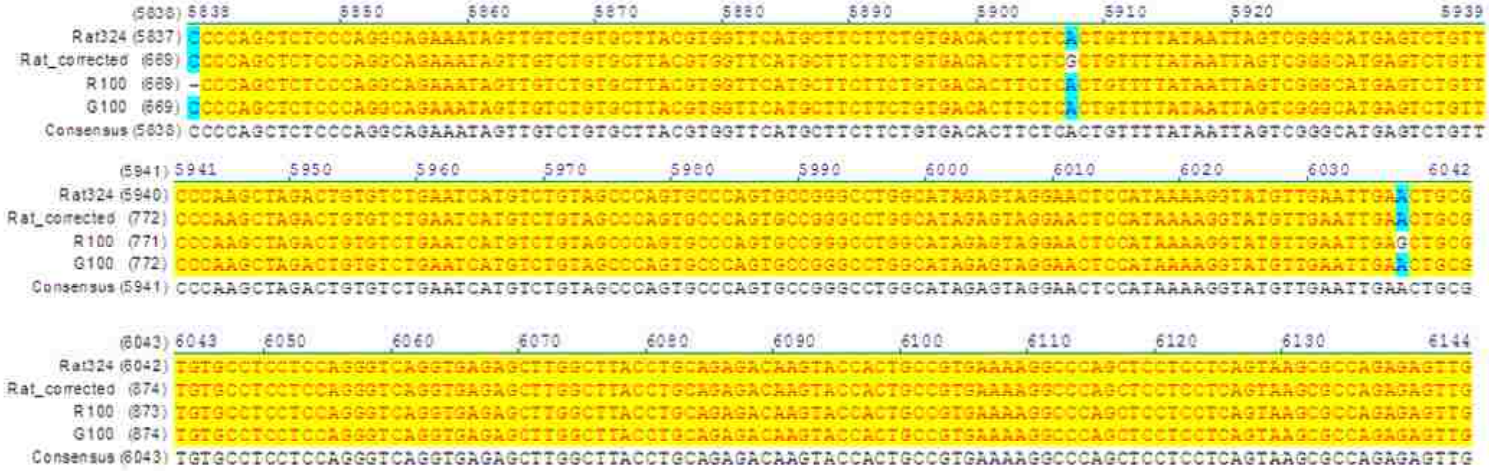
VIII. Appendices

Appendix -A

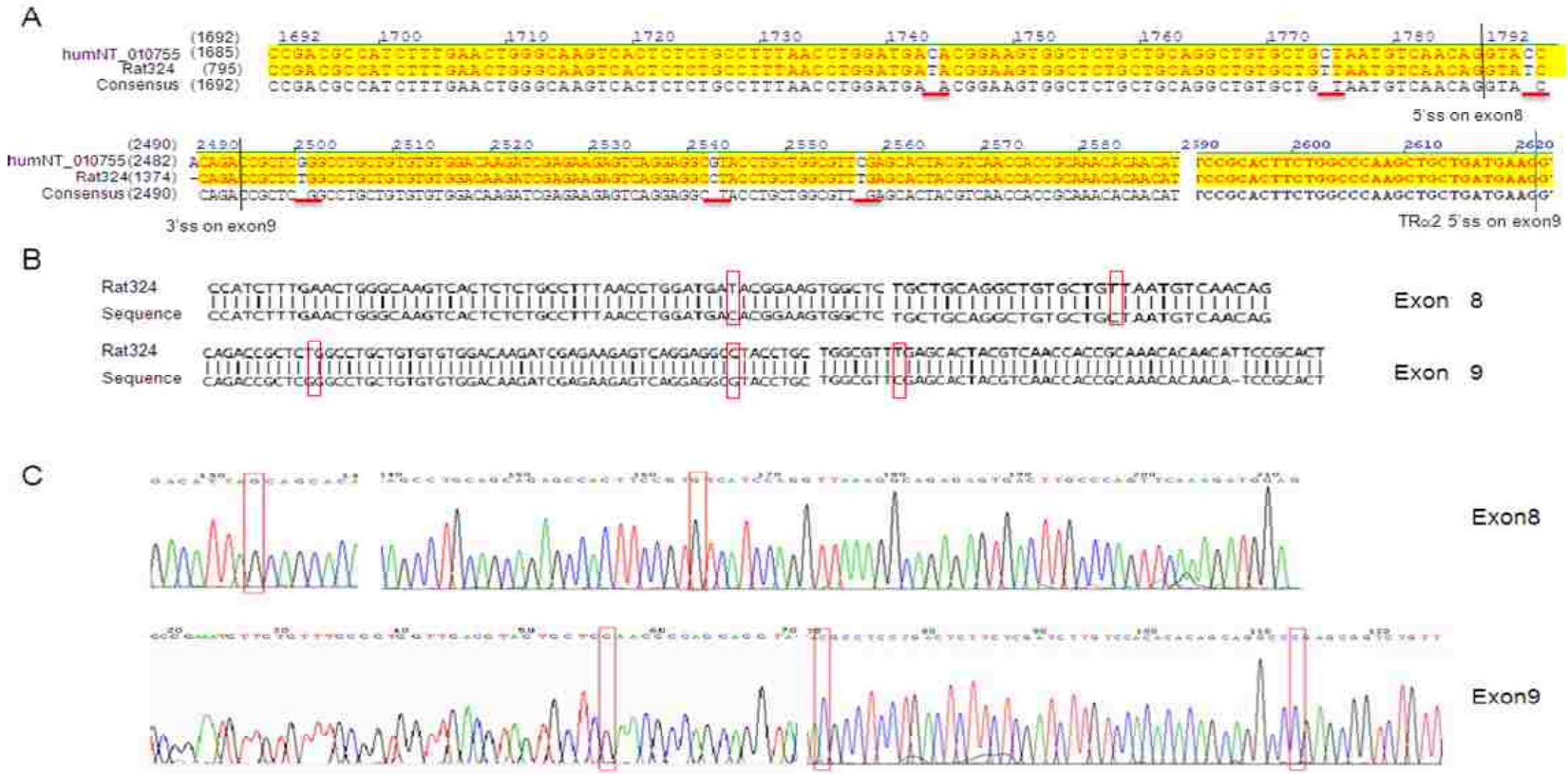
Comparison of authentic rat genome sequence and sequencing results of constructs G100 and R100

Sequence alignment among Genomic rat sequence(Rat324), the corrected sequence from our lab(Rat_corrected), G100 and R100. The boundaries between G100 and R100 are marked. The introduced point mutation in the new made constructs are also indicated in the figure.

Appendix –A(Part II)



Appendix -B



Comparison of partial TRα2 sequence between human and rat species.

A. Authentic sequence alignment for the exon 8 and exon 9A between human (humNT_010755) and rat(Rat324). The red underline marker represents the nucleotides difference in this region. The corresponding splicing sites on exon 8 and exon9 are also indicated.

B. Alignment between isolated TRα2 product from pCMV (Based on Realtime PCR with primer B/B1) and authentic rat TRα2 sequence. The red rectangle circles the different nucleotides from the results.

C. Detailed sequencing graph of the sequence comparison. The single peak in the rectangle indicates that only human TRα2 existed.

Appendix-C

Plasmid name	Catalog #	Description and application	Vector	Application	Source
pHBΔ6	9	β globin sequcne; negative control for in vitro splicing and competition tests	pSP6	in vitro splicing	SHM
pBS500	101	β globin sequcne; negative control for in vitro splicing and competition tests	pGem4	in vitro splicing	SHM
pRC/CMV	210	negative control for transfection and ESX10 study	pCMV	transfection, PCR	SHM
pCMVerbAα	220	wildtype construct for transfection and ESX10 study	pCMV	transfection, PCR	LV
pCMVB324/5'ss	352	disrupted 5'ss of TRa2 and for poly(A) study	pCMV	transfection, PCR	SHM
pα2 ΔBD	493	TRa2 exon8-10; in vitro splicing and competition tests	pGem3	in vitro splicing	SHM
pα2 ΔBE	494	TRa2 exon8-10; in vitro splicing and competition tests	pGem3	in vitro splicing	SHM
pα2SE80ΔBD	505	TRa2 exon8-10 and SEa2 included; in vitro splicing and competition tests	pGem3	in vitro splicing	SHM
pErbA2 ΔBE	555	ΔBE fragment inserted; vector for constructs making	pGem3	in vitro splicing	VS
pErbA2 ΔBE SE60	556	ΔBE fragment inserted and SE60;in vitro splicing and competition tests	pGem3	in vitro splicing	SHM
pDSX100XK	585	exon10 sequnece included; for in vitro transcription	pDsx	in vitro splicing	SHM
pCMV324 α2 pA SVL	613	substitued poly(A) downstream; for poly(A) study	pCMV	transfection, PCR	AS
pErbAMΔ5'ss SVLα2	625	disrupted 5'ss of TRa2 with substituted poly(A) downstream; for poly(A) study	pCMV	transfection, PCR	SHM
pErbAm ΔRM	632	200nt downstream of EcoRI site of exon10 truncated; for poly(A) study	pCMV	transfection, PCR	SHM
pErbAm SVLα2	633	200nt downstream of EcoRI site of exon10 truncated + donwstream poly(A) signal; for poly(A) study	pCMV	transfection, PCR	AS
pErbAm Revβ 200	635	ESX10 replaced by β sequemce; for ESX10 study	pCMV	transfection, PCR	AS
pErbAm Revβ 200 SVLα2	636	ESX10 replaced by β sequemce and introduced poly(A) downstream; for ESX10 study	pCMV	transfection, PCR	AS
pErbAm Revβ G100	647	G100 replaced by β sequemce and introduced poly(A) downstream; for ESX10 study	pCMV	transfection, PCR	CZ
pErbAm Revβ R100	648	R100 replaced by β sequemce and introduced poly(A) downstream; for ESX10 study	pCMV	transfection, PCR	CZ

Appendix-D

Primers for *in vitro* studies

Number	Description of plasmid construct	Direction	Sequence (5' to 3')
701L	pDsx (upstream primer)	F	CGGCCAGTGAATTGTAATACG
			CTAGA GCTCTTCGGG CTCTGGTGCT GAAGAACCGG CCCTCGGAGA CT
705L	Dsx/KX-R42αD	R	TACCGCGATCCAAGCTTATC
			CTAGA GGTGCT GAAGAACCGG CCCTCGGAGA CTTCCCGCTT CACCAA
706L	Dsx/KX-R42αE	R	TACCGCGATCCAAGCTTATC
			CTAGA GG CCCTCGGAGA CTTCCCGCTT CACCAAACTG CTGCTCAAGC
707L	Dsx/KX-R42αF	R	TACCGCGATCCAAGCTTATC
			CTAGA GCACTAAGGA CCTTAATAAT GAAAAACCAT CCAAATGAGG CC
711L	Dsx/KX-R42βD	R	TACCGCGATCCAAGCTTATC
			CTAGA AT CCAAATGAGG CCTCTATTTT TACAAAACTG CTTCTAAAGT
713L	Dsx/KX-R42βF	R	TACCGCGATCCAAGCTTATC
723L	Dsx/KX-R70αFαD'	R	CTAGAGCTCTTCGGGCTCTGGTGCTG
724L	Dsx/KX-R70αFβD'	R	CTAGA GCA CTA AGG ACC TTA ATA ATG AAA AAC CGG CCC TCG GAG ACT TCC CG
725L	Dsx/KX-R70βFαD'	R	CTAGA GCT CTT CGG GCT CTG GTG CTG AAG AAC CAT CCA AAT GAG GCC TCT AT
726L	Dsx/KX-R70βFβD'	R	CTAGA GCA CTA AGG ACC TTA ATA ATG
			CTAGA GGCCCTCGGAGACTTCCCttTTtACCAAACTGCTGCTCAAGC
735L/R	Dsx/KX-R42αFRescue1	R	TACCGCGATCCAAGCTTATC
	42aF(rescue1)	F	GCTTGAGCAGCAGTTTGGTaAAaaGGGAAGTCTCCGAGGGCC
			CTAGA Gca Cta aGG GCT CTG GTG aTG Aaa AAC C GG CCC TCG GAG ACT
736L/R	Dsx/KX-R42αDRescue2,3	R	TACCGCGATCCAAGCTTATC
	42aD(rescue2,3)	F	AGTCTCCGAGGGCC/GGTTtTTCAtCACCAGAGCCCttAGtGC
			GCaCTaaGGGCTCTGGTGaTGAAaAACCGGCCCTCGGAGACT
737L/R	Dsx/KX-R70αFαD'Rescue1,2,3	R	CTAGA GCA CTA AGG ACC TTA ATA ATG AAA AAC C GG CCC TCG GAG ACT TCCCt
			CTAGA GGTGaT GAAaAACCGG CCCTCGGAGA CTTCCCttTTtACCAA
738L/R	Dsx/KX-R42αE(RSCU1,2)	R	TACCGCGATCCAAGCTTATC
720R			GCAGTTTTGTAAAAATAGAGGCCT
	R70aFaD'		GCT TGA GCA GCA GTT TGG TGA AGC GGG AAG TCT CCG AGG GCC /GGT TCT TCA GCA CCA GAG CCC GAA GAG C
	R70bFbD'		ACT TTA GAA GCA GTT TTG TAA AAA TAG AGG CCT CAT TTG GAT /GGT TTT TCA TTA TTA AGG TCC TTA GTG C

759R	Rescue 1,2,3 rectified	R	CTA GA GCAC TAA GGG CTC TGG TGA TGA AAA ACC GGC CCT CGG AGA CTT CCC t - 3'
760R	Rescue1,3-rectified	R	CTAGA GCACTAA GGGCTC TGG TGC TGA AGA ACC GGC CCT CGG AGA CTT CCC t - 3'
761R	R70 aFD(R3)	R	CTA GAG CAC TAA GGG CTC TGG TGC TGA AGA AC
762R	R70 aFD(R1)	R	CTA GAG CTC TTC GGG CTC TGG TGC TGA AGA ACC GGC CCT CGG AGA CTT CCC T
763R	R70 aFD(R2)	R	CTAGA GCTCTTC GGGCTCT GGTGaTG AAaAACC GGCCCTC GGAGACT TC
764R	rescue 2,3	R	CTAGA GCaCTaa GGGCTCT GGTGaTG AAaAACC GGCCCTC GGAGACT TC
765R	rescue 1,2	R	CTAGA GCTCTTC GGGCTCT GGTGaTG AAaAACC GGCCCTC GGAGACT TCCCt
766R	R42xtD	R	CTAGA GCc CTT CGtt CcC T ca TcC TGA AGA Aca Gt Cca aac Gat A Cc TACCGCGATCCAAGCTTATC
767R	R42xtE	R	CTAGA ca TcC TGA AGA Aca Gt Cca aac Gat A Cc TC CCG CTT CAC CAA TACCGCGATCCAAGCTTATC
768R	R42xtF	R	CTAGA Gt Cca aac Gat ACc TC CCG CTT CAC CAA gCT GCT cCT gcg cC TACCGCGATCCAAGCTTATC
769R	R70xtFxtD'	R	CTAGA GCc CTT CGtt CcC T ca TcC TGA AGA Aca Gt Cca aac Gat A Cc TCC CGC
770R	R70xtFaD'	R	CTAGA GC TCT TCG GGC TCT GG TGC TGA AGA ACC Gt Cca aac GAt ACc TCC CGC
771R	R70aFxtD'	R	CTAGA GC cCT TCG ttCcC T ca TcC TGA AGA Aca GG CCC TC GGA GAC TTC CCGCT
773R	Dsx/KX-R42αD (R2)	R	CTAGA GCT CTT CGG GCT CTG GTG aTG AAa AAC C GG CCC TCG GAG ACT TACCGCGATCCAAGCTTATC
774R	Dsx/KX-R42αF (ScR1)	R	CTAGA GG CCCTCGGAGA CTTCCC atccgtc CAAACTG CTGCTCAAGC TAC
775R	Dsx/KX-R42αF (ScR2)	R	CTAGA GG CCCTCGGAGA CTTCCC accctgt CAAACTG CTGCTCAAGC TAC
780R	FD'(R123) primer matched to R100	R	CTAGA GCA CTAA GGG CTC T GG TGA TGA AAA ACC GGC CCT CGG AGA CTT CCC TCG GAG ACT TCC Ctt TTt ACC AAA CTG CTG CTC AAGC
781R	betaG70	R	CTAGA GGA GCT CTT GGC CTT TAA AGT
782R	alphaG70	R	CTAGA GAA GCT GCT GTC CTT CCG GGT
783R	betaG100	R	CTAGA TGC CAG ATC TTC GAT CTT TAA
784R	alphaG100	R	CTAGA TGC CGG ACC TGC GGA CCC TGA
785R	DsxKX	R	CTAGA TA CCG CGA TCC AAG CTT ATC
786F	aG fragment	F	TTC GAT TCT GTA CAA GGG GGC AG
787F	bG fragment	F	TTC GAT TCT GTA CAG TTA CCA TC
788F	Rescue overlap	F	CCC TTA GTG CCG AAG AGC CCG
789R	Rescue overlap	R	CTA GAG ACC TGC TCA ATG CC
790R	Rescue mutated	R	GCA CTA AGG GCT CTG GTG ATG AAA AAC CGG CCC TCG GAG ACT TCC CTC GGA GAC TTC CCT TTT TAC CAA ACT GCT GCC G
791F	recombinant F(paired with 790R)	F	TTT TCA TCA CCA GAG CCC TTA GTG CCC GCA GCA GCG

792F	adding T7 promoter for GR	F	CGA AAT TAA TAC GAC TCA CTA TAG GGA GAT ACA AGG GGG CAG CGG CAG A
793F	adding T7 promoter for bGR	F	CGA AAT TAA TAC GAC TCA CTA TAG GGA GAT ACA GTT ACC ATC AGT TCA T
794F	3'ss probe forward	F	CGA CCT CGA GGG CCC GGT ACC ATT CCA GAG GCT CAT CTT GGA AT
795R	T7 promoter for 3'ss probe	R	CGA AAT TAC GAC TCA CTA TAG GGA GAC CTC GTC AGA GGA AGA GGA G
796F	T7 promoter for alpha R70	F	CGA AAT TAA TAC GAC TCA CTA TAG GGA GAA GCT TGA GCA GCA GTT TG
797R	Reverse for alpha R70	R	GCT CTT CGG GCT CTG GTG CTG
798F	T7 promoter for beta R70	F	CGA AAT TAA TAC GAC TCA CTA TAG GGA GAA CTT TAG AAG CAG TTT TGT A
799R	reverse for beta R70	R	GCA CTA AGG ACC TTA ATA ATG

Primers for *in vivo* studies

Number	Description of plasmid construct	Direction	Sequence (5' to 3')
905R	TRa2 146bp	R	CCT CGT CAG AGG AAG AGG AG
936F	TRa1 skip x9A 291bp	F	CCT GGA TGA TAC GGA AGT GG
937R	TRa1 skip x9A 291bp	R	TCC TCA AAG ACC TCC AGG AA
1467R	rTRa2 135bp	R	TCG AAC TCT GCA CTT CTC TCT C
1468R	rTRa2 140bp	R	CAG AAT CGA ACT CTG CAG TTC T
1469F	rTRa2del5'ss 190bp	F	TGC TGA TGA CCC ACA CAG AT
1470F	rTRa2 5'ss6TG192bp	F	GCT GCT GAT GAA GGT GAC G
1471F	TRa2 wt 5'ss 192bp	F	GCT GCT GAT GGT GAC TG
1472F	CSS1 51bp	F	CGC TTC CTC CAC ATG AGA
1473F	CSS2 48bp	F	GCG GCC AGA GGG AGA
1474F	cross TRa3 185bp	F	ATA CGG AAG TGG CTC TGC TG
1475R	cross TRa3 185bp	R	GAC CTG CGG ACC CTT CAT
1476F	cross TRa3 127bp	F	GCT GTG TGT GGA CAA GAT CG
1477R	cross TRa3 127bp	R	GAC CTG CGG ACC CCT CAT
1478F	cross b200 css 129bp	F	GCT GTG TGT GGA CAA GAT CG
1479R	cross b200 css 129bp	R	CGT GCA CTA AGG ACC TTC ATC
1480F	b200 insert139bp	F	CAC AAC ATT CCG CAC TTC TG
1481R	b200 insert139bp	R	TGA GGA GCT CTT GGC CTT TA
1482R	b200 insert189bp	R	TCA CCC TTA AGG CCT TTG TT
1483F	TRa2 ZCSS-2 89bp	F	TTC CTG GAG GTC TTT GAG GAT
1484R	TRa2 ZCSS-2 89bp	R	CCC CCT TGT ACA GAA TCG AA
1485F	TRa2 ZCSS-2 89bp	F	CCC ACT CTT CCT GGA GGT CT
1490F	R100betaA	F	CCG GCA ACT TTA GAA GCA GTT TTG TA
1491R	G100a B	R	TAA AGT TGC CGG ACC TGC GGA CCC TG
1492F	R100a C	F	CTG GCA GCT TGA GCA GCA GTT TGG TG
1493R	G100betaD	R	TCA AGC TGC CAG ATC TTC GAT CTT TA
900F	Exon 9A	F	TGT GTG TGG ACA AGA TCG AG
1494R	Corrected R123 primer	R	AAA CTG CTG CTC AAG CTG CC
1495R	TRa3 in wt or betaG100	R	TCA CCA AAC TGC TGC TCA A
1496R	TRa3 in betaR100	R	TTT ACA AAA CTG CTT CTA AAG TTG C

1497R	betaR100 or GR200 css	R	TTG CAG GAA ACT CTC ATT CG
1498R	betaR100 or GR200 css	R	AGG CTT TGC AGG AAA CTC TC
1499aR	TRa 9A w/936F 86bp	R	CGA TCT TGT CCA CAC ACA GC
1499bR	TRa 9A w/936F 137bp	R	GGT GGT TGA CGT AGT GCT CA

Primers in realtime PCR analysis

Code	Number	Description of plasmid construct	Direction	Sequence (5' to 3')
A	936F	TRa1 skip x9A 291bp	F	CCT GGA TGA TAC GGA AGT GG
A1	937R	TRa1 skip x9A 291bp	R	TCC TCA AAG ACC TCC AGG AA
A2	1499bR	TRa 9A w/936F 137bp	R	GGT GGT TGA CGT AGT GCT CA
B	900F	Exon 9A	F	TGT GTG TGG ACA AGA TCG AG
B1	1467R	rTRa2 135bp	R	TCG AAC TCT GCA CTT CTC TCT C
B2	905R	TRa2 146bp	R	CCT CGT CAG AGG AAG AGG AG
B3	1495R	TRa3 in wt or betaG100	R	TCA CCA AAC TGC TGC TCA A