REGULATION OF THE 64-kDA SUBUNIT OF CLEAVAGE STIMULATORY FACTOR ACTIVITY IN MACROPHAGE AND B LYMPHOCYTE mRNA 3'-END PROCESSING

by

Scott Allen Shell

BS, Arizona State University, 1999

Submitted to the Graduate Faculty of

The School of Medicine in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2005

UNIVERSITY OF PITTSBURGH

FACULTY OF SCHOOL OF MEDICINE

This dissertation was presented

by

Scott A. Shell

It was defended on

September 14, 2005

and approved by

Martin C. Schmidt, Ph.D.

Sidney M. Morris, Jr., Ph.D.

Scott Plevy, M.D.

Anuradha Ray, Ph.D.

Christine Milcarek, Ph.D.
Dissertation Director

REGULATION OF THE 64-kDA SUBUNIT OF CLEAVAGE STIMULATORY FACTOR ACTIVITY IN MACROPHAGE AND B LYMPHOCYTE mRNA 3'-END PROCESSING

Scott A Shell, PhD

University of Pittsburgh, 2005

Eukaryotic pre-mRNA is processed within the 3'-untranslated region (3'-UTR) resulting in cleavage and polyadenylation. The efficiency of the cleavage reaction is dependent on the binding activity of the 64-kDa subunit of CstF, CstF-64, to the pre-mRNA and is increased with elevated levels of CstF-64. There is evidence that alternative polyadenylation occurs in the presence of increased CstF-64. Our results showed that CstF-64 levels increased with LPS stimulation of RAW 264.7 macrophages while the expression of other pre-mRNA processing factors remained unchanged. Because of the evidence that several macrophage genes exhibit alternative polyadenylation and post-transcriptional regulation under LPS stimulation, we used a reporter mini-gene to identify alternative polyadenylation in LPS-stimulated RAW macrophages. Upon LPS stimulation we measured a 2.5-fold increase in proximal poly(A) site selection that correlated with elevated levels of CstF-64. Forced expression of CstF-64 demonstrated similar alternative polyadenylation. Microarray analysis demonstrated 515 genes changed expression with LPS stimulation, 15 of which also changed with CstF-64 over-expression. A closer analysis of 5 of these 15 genes demonstrated alternative polyadenylation within their 3'-UTR. Closer analysis of the 3'-UTRs showed putative AU-rich regulatory elements. There is also evidence that pre-mRNA processing is coupled with transcription. Previous work has shown that the

carboxy-terminal domain (CTD) of RNAP-II is necessary for 3'-end processing, that CstF binds to RNAP-II CTD and that this binding is CTD phosphorylation dependent. Because our lab has also shown that increases in CstF-64 binding activity upon B-cell differentiation causes alternative polyadenylation on the Ig heavy chain gene and occurs in the absence of CstF-64 increases, we believe that the local concentration of CstF-64 to the nascent pre-mRNA is increased in plasma cells through the phosphorylation-dependent recruitment of CstF-64 to RNAP-II CTD. Using chromatin immunoprecipitation (ChIP), we measured an increase in Serine-2 and Serine-5 phosphorylation of the RNAP-II CTD at the promoter and variable regions of the Ig heavy chain gene in plasma cells compared to memory B cells. We believe that this increase in RNAP-II CTD phosphorylation plays a role in either increased transcription of the Ig heavy chain gene or recruitment of pre-mRNA processing factors to the transcriptional machinery.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

A – Adenosine

AAUAAA – Consensus poly(A) signal

Ab – Antibody

AP – Alkaline Phosphatase

AP1 or AP2 – Adapter Primer specific for the linker used in 5'-RACE PCR

ARE – Adenosine/Uridine Rich Element

AU – Adenosine/Uridine

AUUAAA – Most abundant non-consensus poly(A) signal

AUUUA – Consensus ARE core element

BCR – B Cell Receptor

BMDM – Bone Marrow Derived Macrophage

BSA – Bovine Serum Albumin

C – Cytidine

C – Constant region of Ig

CDK – Cyclin Dependent Kinase

cDNA – Complementary DNA

CF – Cleavage Factor

ChIP – Chromatin Immunoprecipitation

CMV – Cytomegolovirus

CstF – Cleavage stimulatory Factor

CPSF – Cleavage Polyadenylation Specificity Factor

CTD – Carboxy-Terminal Domain

D – Diversity region of Ig

DMEM – Dulbecco's Modified Eagle Medium

DSE – Downstream Element

EST – Expressed Sequence Tag

ELISA – Enzyme Linked Immunosorbent Assay

Eμ - Ig heavy chain intronic enhancer

F – Forward primer used in MRP-PCR

FR1 – PCR product generated by F and R1 primers from MRP-PCR

FR2 – PCR product generated by F and R2 primers from MRP-PCR

G – Guanosine

G/U-rich – Guanosine/Uridine rich

GPCL – University of Pittsburgh Genomics and Proteomics Core Laboratories

Gpt – Guanosyl Phosphotransferase

GRS – Guanosine Rich Sequence

GSP1 or GSP2 – Gene Specific Primer used in 5'-RACE PCR

H5 – RNAP-II CTD phosphorylated serine-2 specific antibody

H14 – RNAP-II CTD phosphorylated serine-5 specific antibody

HRP – Horseradish Peroxidase

hnRNP – Heterogeneous Ribonuclear Protein

Ig – Immunoglobulin

IgA – Alpha form of Ig

IgD – Delta form of Ig

IgG – Gamma form of Ig

IgM – Mu form of Ig

IMDM – Iscove's Modified Dulbecco's Medium

IP – Immunoprecipitation

J – Joint region of Ig

Kb – Kilobase

L – Leader sequence of Ig

LPS – Lipopolysaccharide

MAPK – Mitogen Activated Protein Kinase

mg – milligram

ml – milliliter

mM – millimolar

MRP-PCR – Multiple Reverse Primer PCR

MPA – Mycophenolic Acid

μg – microgram

μl – microliter

ng – nanogram

PABP – Poly(A) Binding Protein

pAm – Membrane poly(A) site of Ig heavy chain gene

PAP – Poly(A) Polymerase

pAs – Secretory poly(A) site of Ig heavy chain gene

PBS – Phosphate Buffered Saline

PBS/T – PBS with 1.0% Tween-20

PC – Plasma Cell

PCR – Polymerase Chain Reaction

P/D – calculated ratio of Proximal to Distal protected fragments from RNase protection assay

PI – Protease Inhibitors

p-NNP – p-Nitrophenyl Phosphate

QPCR – Quantitative or real-time PCR

R1 – Proximal reverse primer used in MRP-PCR. Measures the abundance of all transcripts of a gene regardless of size.

R2 – Distal reverse primer used in MRP-PCR. Measures the abundance of transcripts processed at the distal poly(A) site.

RNAP II – RNA Polymerase subunit II

RRM – RNA Recognition Motif

RT-PCR – Reverse Transcriptase PCR

SEM – Standard Error of the Mean

SS – Splice Site

T – Thymidine

TBP – TATA Binding Protein

TCA – Trichloro Acetic Acid

TF – Transcription factor

TLR – Toll Like Receptor

TNF-α - Tumor Necrosis Factor alpha

U – Uridine

UTR – Untranslated Region

USE – Upstream Element

V – Variable region of Ig

VDJ – Recombined Variable, Diversity and Joint segments of the Ig heavy chain gene

PREFACE

The work contained within these pages has been quite a task to accumulate. Although the work described herein is close to 100% my own, I have many people that deserve praise for supporting me through these endeavors. First and foremost, I thank my mentor Christine Milcarek, who has provided me and my project with financial support, and has also provided me with a degree of freedom that allowed me to take these projects into directions never originally anticipated; directions that have allowed us to demonstrate the importance of our work to the scientific community. By working in her lab, I have not only gained experience in powerful contemporary techniques, but I have also presented my work at international scientific meetings and have been charged with writing my own manuscript for publication. I thank the members of my thesis committee for their intellectual contribution and support of my thesis, and in particular Dr. Sidney M. Morris, Jr. for his contribution to our manuscript in which he appears as co-author. Patrick Grof-Tisza and Candice Hesse were two undergraduate students that I had mentored over two consecutive summers. By interning in the lab, they had provided me with an opportunity to teach scientific principles and laboratory techniques that had not only widened my breadth of experience in graduate school, but had also forced Dr. Milcarek and me to come up with projects that, originally not planned for immediate execution, ended up significantly contributing to my body of work.

I thank Debbie Hollingshead with the GPCL for her continued help and advice on the experimental design for the QPCR and microarray experiments that we performed in their facilities, Robert Baron who had designed the computer algorithm that demonstrated our statistical strength of our overlap data from the microarray experiments, and Dr. Joann Flynn and

her lab for supplying us with the mouse bone-marrow macrophages we used to demonstrate CstF-64 protein increases with LPS stimulation.

One of my fondest memories of interviewing for the IBGP was my fiery and energetic interview with Dr. Will Walker in Cell Biology and Molecular Physiology. I subsequently spent my second rotation under his tutelage. He and his then graduate student, Dr. Frank Delfino, had both helped me a great deal in training me as a bench scientist, for I was quite green upon entering graduate school despite my one year of work in a biochemistry lab at Arizona State University my senior year. Dr. Walker taught me the value of having a vision when it comes to mapping out a project. The first day I was in his lab, he sketched out on a legal pad what he wanted me to accomplish during my rotation. Subsequently, while working by my side in the lab, I was able to achieve all he had set out for me to accomplish resulting in my rotation project appearing in my premier first-authorship manuscript. Frank continues to be my scientific colleague, molecular biology go-to guy and personal friend and would still whip me in racquetball if I dare step into a court with him again.

I also thank a great bunch of guys that I had met through the graduate program. I met Dominic Warrino during our first summer rotation and we hit it off instantly with our common interests in card playing, beer drinking and listening to metal music. It sounds like a friendship doomed to a run in with the authorities, however, toward the end of his graduate career, he and his wife ended up spending considerable time with me and my family picnicking, BBQing and eating the habenero hot wings at Fat Heads on the South Side to a point where we couldn't see. Chris Herrem and I have become quite good friends over the years. Not only have we had some great laughs together, we have been there for each other during some of the more difficult times graduate school can throw at you. With our common interests in politics, military history and

NTN trivia at Damon's, it is no wonder that we still are quite close friends. I met Andy Lepisto the same day I met Chris Herrem. I have had the fortunate opportunity to find someone that has the same penchant for cigar smoking, Texas Hold 'Em and horseshoes as I do. Apparently we were both cowboys in a previous life. He also has been a good sounding board when discussing future endeavors that lie beyond graduate school. I cannot remember when I met Kamal Khanna. However, he quickly became one of my most intellectual peers where we would have discussions lasting hours regarding science, religion and international politics and history. Usually, the length of our conversations were directly proportional to the amount of scotch consumed, and I have to say that since he has left for his post-doctoral pursuits, there has been a void in the intellectual element of my social life. Mike Turner will be graduating soon. He and I are two of the last few that will be graduating from our class and that, in and of itself, provides an instant kinship. Our projects are similar only in that there were unforeseen hang-ups in our work which caused significant delays that held each of us up for several months. Because misery loves company, these events have bonded our friendship in ways hashing through the streets of Squirrel Hill on a full moon never could. All of the above mentioned gentlemen have demonstrated dedication and persistence in the pursuit of scientific excellence and their graduate degree and all have been and continue to be an inspiration to me.

Lastly, I thank the members of my family. My mother has been a pillar of strength for me from the day I left for Arizona State University many years ago and has been there for me emotionally and, sorry to say, financially ever since I embarked on my graduate school journey. With her unconditional support only an outstanding mother could provide, I have been able to keep focused on my goals which resulted in me obtaining a graduate degree from a top research institution of which I am very proud to have earned. I only hope to feel the pride in myself that I

have bestowed upon her in accomplishing the feat of earning my Ph.D. I thank my brother for his continued interest in my progress through the years. As my best friend, I anticipate many years of most excellent times raising our families side-by-side in Chicago (and maybe forming a rock cover band in the very distant future). My wife, Jenifer, has been my focal point for the ten years we have known each other. She has always been supportive in all my professional endeavors. Testament to that is her willingness to leave everything and everyone she knew to travel across the country to a strange place without a job to be with me during my academic and professional pursuits. I cannot say that things have been easy emotionally or financially, but six years and two wonderful children later (Jackson and Jamison), we have finally made it. She has continually been there for me, particularly during the prototypical third-year blues many of us go through wondering why in the hell we have chosen the path we have. After mustering a fighting spirit, here we are on the tail end of this sojourn. Our immediate future is still being worked out, but the idea of what we are wanting for us and our children is alive and becoming less abstract and more corporeal. What an exciting time! Lastly, I would like to pay homage to my late father, Jerome Q. Shell. He died of pancreatic cancer several years after I left for Arizona and I hadn't the opportunity to spend a great deal of time with him while I was away. He died before I graduated with my B.S. and had never learned of my desire to go beyond earning my undergraduate degree. Regardless of his being absent for the last ten years of my life, I would not have been able to do any of this without him. Apparently his tenacity and continual pursuit of education and self-improvement while he was alive has manifested itself in me. Because of that, I dedicate my Ph.D. to him. Serendipitously, my beautiful and sparky daughter, Jamison, was born on my late father's birthday.

1. STATEMENT OF THE PROBLEM

The importance of alternative polyadenylation to gene expression has been demonstrated for the Immunoglobulin (Ig) gene. Selection of the promoter-distal poly(A) site on the Ig gene allows inclusion of membrane spanning exons into the final gene transcript, thus creating the membrane form of Ig that is primarily expressed in naïve and memory B lymphocytes. Upon B-cell differentiation into a plasma cell, there is a switch to the use of the promoter-proximal poly(A) site that precludes the incorporation of the membrane spanning exons, thus creating the secretory form of Ig. Alternative polyadenylation demonstrates influences on gene expression beyond the change in the final structure of the protein as shown for Ig. For example, alternative polyadenylation can influence gene expression when a change in poly(A) site use precludes the incorporation of mRNA regulatory elements within the 3'-UTR that control stability and translation of the final gene transcript. This mechanism of gene expression control has been demonstrated in a variety of genes of the immune system that are involved in the acute defense response such as cytokines, COX-2, lysozyme and TNF-α. Although the mechanism of cleavage and polyadenylation is well understood, the regulation of selection of one poly(A) site over another is less clear and appears to be influenced by a variety of things including, but not limited to: 1) competition between splicing and 3'-end processing, as occurs on the Ig gene; 2) the abundance of cleavage and polyadenylation factors, in particularly CstF-64, in the cellular environment; 3) the strength of one poly(A) site over another, which is strongly influenced by the exact sequence of and around the cleavage and polyadenylation site; and 4) the abundance of auxiliary factors that regulate the cleavage and polyadenylation machinery by either increasing or decreasing its efficiency.

Although alternative polyadenylation events have been described on a number of LPS-induced macrophage genes, the cause of these events has not been elucidated. Through the studies outlined in this thesis, we have attempted to elucidate the contribution CstF-64 has on both the alternative polyadenylation of genes that possess more than one poly(A) site and on global gene expression in general in macrophages activated by gram-negative bacteria.

Furthermore, there has been substantial evidence linking transcription and pre-mRNA 3'-end processing on yeast genes by demonstrating that yeast mRNA processing factors are recruited to the CTD of RNAP-II during transcription in a phosphorylation dependent manner. Because similar work in the mammalian system has yet to be performed, particularly on an inducible gene, and because of the impact 3'-end processing has on the expression of Ig heavy chain mRNA isoforms, we reasoned that the phosphorylation state of RNAP-II CTD across the Ig heavy chain gene locus may be different between memory B-cells and Ig-secreting plasma cells. We also reasoned that these differences in the phosphorylation state of RNAP-II CTD may be associated with changes in mammalian mRNA processing factor binding to the transcriptional machinery, thus influencing poly(A) site switch on the Ig heavy chain gene transcript. Therefore, we have attempted to demonstrate a link between transcription and pre-mRNA cleavage and polyadenylation of the Ig heavy chain gene by identifying changes in the phosphorylation state of RNAP-II CTD across the Ig heavy chain genomic locus that correlates with the increased production of the secretory form of Ig that is observed when B-cells differentiate into Ig-secreting plasma cells.

2. SPECIFIC AIMS

AIM 1 – We will test the hypothesis that the stimulation of the RAW 264.7 murine macrophage cell line with LPS or IFN-γ causes an increase in 3'-end processing activity that correlates with an increase in CstF-64 protein expression.

AIM 1a. We plan to assess by Western blot changes in protein expression for an array of factors and enzymes involved in mRNA 3'-end processing between unstimulated, LPS-stimulated, and IFN-γ–stimulated RAW macrophages. We also plan to measure the same protein expression profile in murine, bone marrow macrophages (BMMs) stimulated with LPS or INF-γ.

AIM 1b. We plan to measure changes in 3'-end processing activity between untreated and LPS or IFN-γ treated RAW macrophages. We aim to correlate any measured change in 3'-end processing activity after stimulation to changes in the protein expression profile of the factors monitored in AIM 1a that occur under the same stimulating conditions.

AIM 1c. We plan to over-express recombinant CstF-64 in RAW 264.7 macrophages and measure changes in 3'-end processing activity in the absence of LPS stimulation.

AIM 1d. We plan to perform microarray analysis on untreated, LPS-stimulated and CstF-64 over-expressing RAW macrophages in an attempt to identify similar changes in gene expression between the LPS-stimulated and CstF-64 over-expressing RAW macrophages.

AIM 1e. We plan to identify alternative polyadenylation events within the 3'-UTR of genes in RAW macrophages whose expression has changed under both LPS-stimulating and CstF-64 over-expressing conditions.

AIM 2 – We will test the hypothesis that differential phosphorylation of RNAP II CTD along the Ig gene locus occurs between B-cells and plasma cells. By employing the technique of chromatin

immunoprecipitation (ChIP) along the length of the Ig gene, we aim to ascertain the phosphorylation state of the CTD of RNAP II along the length of the Ig heavy chain gene. Any changes in the phosphorylation state of the CTD upon B-cell differentiation will be determined by comparing ChIP results between A20 memory B-cell and AxJ plasma cell lines which share the IgG2a gene but regulate poly(A) site usage on the IgG2a pre-mRNA differently.

3. INTRODUCTION

3.1. mRNA Processing

The amount of protein product that results from gene activation is influenced by multiple biological mechanisms, including the rates of transcription, mRNA processing, mRNA export to the cytoplasm, mRNA stability and translational efficiency of the gene transcript. Post-transcriptional RNA processing is essential in creating translationally competent mRNA. The 5'-end of pre-mRNA is capped with a modified guanosine residue placed in a 5'-5' orientation. Non-coding intron sequences are removed by splicing. Endonucleolytic cleavage of nascent pre-mRNA occurs primarily at an adenosine residue within its 3'-UTR, although it has been determined that endonucleolytic cleavage can occur at alternative residues resulting in minor heterogeneity at the extreme 3'-end of the transcript (106). Following 3'-end cleavage is the addition of ~200 adenosine residues, thus creating a poly(A) tail [reviewed in (43) and (158)]. Proteins, such as poly(A) binding protein (PABP) and eIF4E, that possess specific poly(A) tail and 5'-cap binding capabilities, respectively, promote interaction between the two ends of the mature mRNA conferring stability, exportability, and translational efficiency of the transcript [reviewed in (150)].

3.2. Required Factors for mRNA 3'-end Processing

In vitro reconstitution studies have revealed that there are a minimum number of defined factors necessary for successful pre-mRNA 3'-end cleavage and polyadenylation. These factors include the Cleavage Polyadenylation Specificity Factor (CPSF) (see below) that binds the conserved AAUAAA poly(A) signal and the Cleavage Stimulatory Factor (CstF) (see below) that binds to the downstream GU-rich region (Figure 1). Other required factors include

mammalian Cleavage Factor I and II (CF-I_m and CF-II_m), poly(A) polymerase (PAP) and poly(A) binding protein (PABP) [reviewed in (43) and (158)], and are not discussed further here.

3.2.1. Cleavage Polyadenylation Specificity Factor

The Cleavage Polyadenylation Specificity Factor (CPSF) is composed of four protein subunits: a 30-kDA subunit (CPSF-30), a 73-kDa subunit (CPSF-73), a 100-kDa subunit (CPSF-100) and a 160-kDa subunit (CPSF-160) (11). CPSF-160 contains an RNA recognition motif (RRM), divergent from other conserved RRMs (96), that binds to the consensus poly(A) signal AAUAAA (71) found in the majority of RNA polymerase II transcribed eukaryotic genes (Figure 1). Protein-protein interaction between the CPSF complex and the CstF complex is facilitated through CPSF-160 (96) and is required for efficient pre-mRNA 3'-end cleavage and polyadenylation. Recently, CPSF-73 has been described as a member of the metallo-β-lactamase family of enzymes (4) and implicated as the 3'-end processing factor that is responsible for the endonucleolytic cleavage of the pre-mRNA (123). UV-crosslinking studies have shown that CPSF-73 binds to the pre-mRNA near the cleavage site in an AAUAAA-dependent manner and the removal of Zn^{+2} from the *in vitro* reaction demonstrated a requirement for metal ions for the cleavage reaction, a requirement other members of the metallo-β-lactamase enzyme family exhibit (123).

3.2.2. Cleavage Stimulatory Factor

The Cleavage Stimulatory Factor (CstF) complex is composed of three protein subunits: a 50-kDa subunit (CstF-50), a 64-kDa subunit (CstF-64), and a 77-kDa subunit (CstF-77). Homologues of CstF-64 and CstF-77 have been described in *Saccharomyces cerevisiae* (37, 57) *Drosophila melanogaster* (61, 94) and *Bos Taurus* (34). The CstF trimeric complex is assembled by CstF-77 forming a bridge between CstF-64 and CstF-50 (135). CstF-77 is also the subunit

that tethers CstF to the 160-kDa subunit of the CPSF (96) and contains a nuclear localization signal that allows transport of the CstF trimer into the nucleus where CstF functions (135). CstF-50 possesses seven contiguous WD-40 protein-protein interaction domains (137) that are necessary for binding to CstF-77 (136). Furthermore, the WD-40 repeats appear to play a role in the binding of CstF-50 to the CTD of RNA polymerase II (89) and to BRCA1-associated RING domain protein (BARD1) (74), and may be responsible for regulation of CstF function. CstF-64 is comprised of a conserved RNA recognition motif (RRM) near its N-terminus (134) that binds preferentially to GU-rich regions [(138) and Figure 1]. Structural analysis of CstF-64 by nuclear magnetic resonance has shown that the RRM forms a β-sheet and demonstrates highest affinity to GU-rich regions that contain a pair of consecutive uridines (108). CstF-64 also contains a hinge region immediately downstream of the RRM that is responsible for binding to CstF-77 and symplekin (136), a protein known to also bind CPSF (66). Two proline/glycine-rich domains surround twelve repeats of the consensus MEARA/G amino acid sequence near the C-terminus that form an alpha-helix secondary structure (134) that has been ascribed as having protein binding characteristics similar to alpha-helical possessing transcription factors (115). The function of the C-terminal domain of CstF-64 has yet to be described, but a report that demonstrates an interaction between the C-terminal domain of CstF-64 and the transcription elongation factor PC4 (18) further supports the view that transcription and pre-mRNA processing are physically linked mechanisms (32, 47, 81, 124).

3.2.2.1. The Variant Form of CstF-64

It has been shown that many mRNAs from male germ cells do not contain the canonical poly(A) signal AAUAAA, yet are efficiently processed to mature mRNA (147). Although there is approximately 250-fold more CstF-64 mRNA expressed in testicular cells compared to liver

cells (33), further investigation of the known CstF-64 in mouse male germ cells revealed absence of its expression in mitotic spermatocytes. This was explained through the mapping of the CstF-64 gene to chromosome X (147), a chromosome inactivated during male meiosis (60). However, a 70-kDa variant form of CstF-64 was found to be expressed in mitotic spermatocytes and post-mitotic spermatids (147). The mouse variant form of CstF-64, called mτCstF-64, was mapped to chromosome 19 and explains how mRNAs in mitotic germ cells can be properly processed in the absence of expression of the X-chromosomal form of CstF-64 (35). A human form of the CstF-64 variant, hτCstF-64, was similarly mapped to chromosome 10 (34). A critical difference between the somatic CstF-64 and τCstF-64 is a Proline-to-Serine amino acid substitution in the RRM, which may instill altered RNA binding specificity, thereby explaining the differences in cleavage and polyadenylation observed in male germ cells (35). Other amino acid substitutions between both the mouse and human forms of CstF-64 and τCstF-64 are the result of the exclusion of protein coding exons (34). Critical analysis of the functional differences between these alternative forms of CstF-64 has yet to be performed. Interestingly, a reciprocal translocation at 10q22 observed in an oligospermic male was mapped to the location of the hτCstF-64 gene (14), supporting the hypothesis that lesions in this gene, although not fatal, may cause male sterility.

3.2.3. Auxiliary Factors that Influence pre-mRNA 3'-end Processing

The heterogeneous nuclear ribonucleoproteins H, H' and F, three proteins that have similar sequence identities and possess RRM-like regions (67), have been implicated in the regulation of pre-mRNA splicing (50) and 3'-end processing (6, 68). Studies from our lab have shown that hnRNP H' and hnRNP F have some influence on the regulated cleavage and polyadenylation on the Ig heavy chain gene in B-lymphocytes (146). Specifically, the elevated

expression of hnRNP F relative to hnRNP H' observed in memory B-cells reduced the binding activity of CstF-64 to GU-rich regions downstream of the Ig secretory poly(A) site (pAs) (146). However, reduced expression of hnRNP F in Ig-secreting plasma cells lifts its inhibitor effect on CstF-64 binding to the pAs allowing for an increase in secretory Ig production (146). Furthermore, we have shown that the transcriptional elongation factors PC4 and ELL2 are upregulated in plasma cell (data not shown). PC4 and ELL2 are proteins that regulate RNAP-II transcription elongation rates [(51, 131) and reviewed in (130)]. Furthermore, PC4 has been shown to interact with the C-terminal domain of CstF-64 (18) and elongation factors have been shown to regulate the incorporation of skipped exons in *in vivo* splicing reactions (70). Because of the implicated balance of splicing and cleavage/polyadenylation that occurs at the secretory poly(A) site of the Ig gene during B-cell development (15, 110), the roles PCR and ELL2 specifically play in the selection of pAs in plasma cells is currently being investigated in our lab.

3.3. Factors that Contribute to Poly(A) Site Strength

The process of post-transcriptional pre-mRNA cleavage and polyadenylation, although essential in creating translationally competent mRNA, is a greatly under-appreciated contributor to the regulation of gene expression. It has been estimated that only approximately 30% of all pre-mRNA transcripts made are successfully processed into a mature RNA species (69). The frequency with which a poly(A) site is selected for the cleavage/polyadenylation reaction is greatly impacted by the strength of that poly(A) site. One factor in *cis* that contributes to the strength of a poly(A) site is the deviation of the poly(A) signal from the consensus sequence AAUAAA (129, 151). In fact, database analysis has shown that as many as 26% of all human genes possess poly(A) signals that deviate from the consensus AAUAAA (9) and that non-consensus poly(A) signals are less often selected for 3'-end cleavage. One study has discovered

that a naturally occurring mutation in the poly(A) signal of alpha-2 globin to AAUAAG results in a significantly reduced accumulation of the gene product (64). Other factors that contribute to the strength of a particular poly(A) site are the U-richness of upstream elements (USE) and downstream elements (DSE) that flank the poly(A) site (79, 98), and the distances the poly(A) signal and the GU-rich region are from the cleavage site (21). More specifically, Chen *et al*. determined that cleavage cannot occur closer than 11 nor farther than 23 nucleotides from the poly(A) signal (21). Furthermore, the GU-rich region is usually located between 10 and 30 nucleotides downstream from the cleavage site [(21) and Figure 1A]. The efficiency of 3'-end processing drops dramatically when the distances of these *cis*-elements are located outside the ranges listed above. An additional G-rich sequence (GRS) has been identified downstream of the CstF-64-binding GU-rich region. Binding of the *trans*-acting factor hnRNP H' to the GRS has demonstrated its contribution to the efficiency of the cleavage reaction (7). Secondary structures of the 3'-UTR have also been implicated in contributing to the strength of viral poly(A) sites (152). The frequency of cleavage/polyadenylation at a specific poly(A) site is also sensitive to the abundance and/or activity of *trans*-acting mRNA processing factors in the cellular environment (24, 139, 146).

3.4. Types of 3'-end regulation

3.4.1. Differences in poly(A) tail length and shortening.

An additional regulatory influence the 3'-UTR has on gene expression pertains to the controlled length of the poly(A) tail. The mouse form of tumor necrosis factor-α (TNF-α) mRNA demonstrates heterogeneous sizes between unstimulated and LPS-stimulated macrophages (31). Closer analysis revealed that this heterogeneity results from changes in poly(A) tail length, where mRNA in unstimulated macrophages possesses an uncharacteristically

short poly(A) tail. Subsequent stimulation of the macrophages with LPS causes a marked increase in length of the poly(A) tail, resulting in enhanced mRNA stability, mRNA translation and protein production (31). Similarly, an increase in the poly(A) tail length of the lysozyme gene transcript occurs as a result of LPS stimulation of the HD11 myelomonocytic cell line and significantly contributes to the accumulation of lysozyme mRNA (54).

3.4.2. AU-rich Elements and Their Contribution to Gene Expression

The stability of a particular mRNA species can have a profound effect on the amount of protein product generated. Excluding the contributions of transcription, the mere increase in the half-life of an mRNA has been shown to directly correlate with the amount of protein produced. *Cis*-elements that control mRNA half-life are typically AU-rich and often possess a consensus AUUUA core element [reviewed in (150)]. However, novel non-AUUUA containing AREs with looser sequence requirements have been identified (20, 107). An abundance of *trans*-acting factors have been shown to bind to AREs and exhibit a variety of functions. For example, binding of AUF1 to the monomer UUAUUUAUU promotes mRNA instability (36), while binding of HuR/HuA to AUUUA-containing mRNA elements increased the half-life of the mRNA (46, 97). Interestingly, it has been shown that the TIA-1 and TIAR proteins are recruited to AREs and regulate translation on TNF-α (59, 111). COX-2 is a gene whose 3'-UTR possesses both mRNA stability elements and translational control elements (27) and demonstrates alternative polyadenylation under various cellular stimuli (101).

3.4.3. Alternative Polyadenylation

As the majority of eukaryotic RNAP-II transcribed mRNAs possess a poly(A) site that facilitates the addition of a poly(A) tail, scores of genes have been identified that possess more than one poly(A) site [reviewed in (45) and (140)]. The choice of one poly(A) site over another

can have significant consequences on the expression and, in some cases, the structure of the final gene product (45).

3.4.3.1. Consequences of Alternative polyadenylation

Database analysis of human ESTs has revealed that approximately 54% of human genes and 32% of mouse genes possess multiple poly(A) sites (140) and that selection of one poly(A) site over another is tissue-specific (10). In general, selection of one poly(A) site over another on the same gene will result in an alteration of the 3'-end of the mature transcript. According to the review by Edwalds-Gilbert *et al.*, there are three categories of exon arrangement for a gene with multiple poly(A) sites [(45) and Figure 2]. The first arrangement type is called "Tandem Poly(A) Sites" and occurs when all poly(A) sites of a gene are located downstream from the last protein coding exon (Figure 2A). In this case, alternative polyadenylation will alter the length of the 3'-UTR, but it will have no effect on the structure of the protein. However, choice of one poly(A) site over another can determine the absence or presence of regulatory *cis*-elements (AREs) in the final transcript, thereby influencing mRNA half-life and translation. The human and murine COX-2 genes are well-studied examples of how alternative polyadenylation can profoundly effect the expression of the gene by the elimination or retention of mRNA regulatory elements (26, 27, 101).

The second arrangement is called the "Composite Internal/terminal Exon" and occurs when a poly(A) site lies between two protein coding exons (Figure 2B). Use of the more upstream poly(A) site will result in a shorter mature gene transcript that excludes downstream protein coding regions. This would result in a C-terminal truncated form of the protein. However, skipping of the promoter-proximal poly(A) site occurs from the selection of an internal splice site that lies within the exon immediately upstream from the proximal poly(A) site.

Selection of this upstream internal splice site will subsequently create a composite exon of the one containing the internal splice site and the downstream exon possessing the splice site acceptor. This splicing event removes the proximal poly(A) site from the mature transcript along with the intron it lies within creating a composite exon and a larger form of the protein. The most studied example of this form of alternative polyadenylation is the immunoglobulin gene in B-lymphocytes (see below). Another example of the "Composite Internal/terminal Exon" alternative polyadenylation event is the NF-ATc transcription factor (25). Selection of the distal poly(A) site occurs in naïve T-cells and creates a longer form of NF-ATc that exhibits reduced trans-activating effects. In effector T-cells, the shorter form of NF-ATc is produced as a result of the increased selection of the promoter-proximal poly(A) site (25). Another rather unique example occurs on the gene transcript for the FCA RNA-binding protein in *Arabidopsis* (82). FCA auto-regulates its expression when the full-length version of the protein (encoded by 13 exons) binds, with the pre-mRNA 3'-end processing factor FY, to it's own transcript and promotes the selection of the proximal poly(A) site (132). This proximal poly(A) site selection results in cleavage and polyadenylation after the third exon of the transcript and produces a non-functional truncated form of the protein. The combination and timing of these two forms of FCA exhibit flowering control in *Arabidopsis* (116).

The third and final arrangement is called "Skipped Exons" and occurs on genes that contain two or more terminal exons each with their own downstream poly(A) site (Figure 2C). An alternative splicing event would then control which C-terminal exon is incorporated into the final gene product. Therefore, two proteins with alternative C-terminal domains can be produced. One example of this is the sex-determining doublesex (dsx) *Drosophila melanogaster* gene (63). Splicing of exons 1, 2, 3, 5 and 6 result in the use of the exon-6 associated poly(A)

site and results in a male phenotype. Conversely, splicing of exons 1, 2, 3, and 4 result in the use of the exon-4 associated poly(A) site and results in a female phenotype (17).

4. AIM 1 AND THE FIRST MANUSCRIPT FOR PUBLICATION: Elevated Levels of the 64-kDa Cleavage Stimulatory Factor (CstF-64) in LPS-Stimulated Macrophages Influence Gene Expression and Induce Alternative Poly(A) Site Selection (Long Form)

4.1. ABSTRACT

LPS activation of murine RAW 264.7 macrophages influences the expression of multiple genes through transcriptional and post-transcriptional mechanisms. We observed a 5-fold increase in CstF-64 expression following LPS treatment of RAW macrophages. The increase in CstF-64 protein was specific in that several other factors involved in 3'-end processing were not affected by LPS stimulation. Activation of RAW macrophages with LPS caused an increase in proximal poly(A) site selection within a reporter mini-gene containing two linked poly(A) sites that occurred concomitant with the increase in CstF-64 expression. Furthermore, forced over-expression of CstF-64 protein also induced alternative poly(A) site selection on the reporter mini-gene. Microarray analysis performed on CstF-64 over-expressing RAW macrophages revealed that elevated levels of CstF-64 altered the expression of 52 genes, 15 of which showed similar changes in gene expression with LPS stimulation. Sequence analysis of the 3'-untranslated regions of these 52 genes revealed that over 46% possess multiple putative poly(A) sites. Five of these 52 genes demonstrated alternative polyadenylation under both LPS-stimulating and CstF-64 over-expressing conditions. We conclude that the physiologically increased levels of CstF-64 observed in LPS-stimulated RAW macrophages contribute to the

changes in expression and alternative polyadenylation of a number of genes, thus identifying another level of gene regulation that occurs in macrophages activated with LPS.

4.2. INTRODUCTION

As a critical component of the innate immune system, macrophages respond to LPS from gram-negative bacteria by activating multiple signaling pathways that result in the rapid expression of pro-inflammatory cytokines and arachidonic acid metabolites. The induced expression of many of these genes is attributed to the activation of transcription factors linked to Toll-like Receptor 4 (TLR4) related signal transduction pathways [reviewed in (104) and (143)]. Aside from transcriptional contributions to gene expression in stimulated macrophages, post-transcriptional mechanisms have also been shown to influence gene expression. For example, the half-lives of TNF-α (31) and lysozyme (54) mRNAs are increased following LPS-stimulation by a lengthening of the poly(A) tail. Additionally, translational regulatory and/or mRNA stability elements have been identified in the 3'-untranslated region (3'-UTR) of human Il-1ra (153) and the LPS-inducible mRNAs encoding COX-2 (26, 27, 38) and TNF-α (76, 157).

Post-transcriptional pre-mRNA cleavage and polyadenylation, although essential in creating translationally competent mRNA, is a greatly under-appreciated contributor to the regulation of gene expression. Because it has been estimated that only 30% of primary transcripts are polyadenylated (69), changes in 3'-end processing can have a profound effect on the amount of mature transcript produced. The frequency with which a poly(A) site is selected for the cleavage/polyadenylation reaction is greatly impacted by the strength of that poly(A) site. Relevant to the present study, the frequency of cleavage/polyadenylation at a specific poly(A)

site is also sensitive to the abundance and/or activity of *trans*-acting mRNA processing factors (24, 139, 146).

Studies of pre-mRNA processing *in vitro* have identified soluble *trans*-acting factors that are required for the cleavage and polyadenylation of gene transcripts. The 160-kDa subunit of the tetrameric complex Cleavage Polyadenylation Specificity Factor (CPSF-160) binds to the consensus poly(A) signal AAUAAA that is located upstream of most functional poly(A) sites (96). The Cleavage Stimulatory Factor (CstF) trimeric complex binds to the downstream element, a guanosine/uridine rich (GU-rich) region located downstream of the cleavage site (21) through its 64-kDa subunit (CstF-64) (138). CstF-64 is comprised of a conserved RNA recognition motif (RRM) near its N-terminus (134) that binds preferentially to GU-rich regions (138). Additional factors involved in pre-mRNA processing include Cleavage Factor I_m and II_m and poly(A) polymerase (PAP). We investigated the expression of CstF-64 and other cleavage/polyadenylation factors after stimulation of murine RAW 264.7 macrophages and murine bone marrow-derived macrophages (BMDMs) with LPS. We found that LPS stimulation of RAW macrophages and murine BMDMs for 18 hours significantly increased CstF-64 protein expression. Because LPS stimulation of macrophages arrests the cell cycle (144, 145), the increase in CstF-64 protein expression occurs in the absence of cellular proliferation. This is in contrast to the increases in CstF-64 expression observed in cells that were induced to proliferate (86, 139). Because exogenous over-expression of CstF-64 in chicken B-lymphoma DT40 cells increased the use of the weak poly(A) site in the IgM heavy chain gene product (139), we hypothesized that the increased levels of CstF-64 observed in LPS-stimulated macrophages would have a similar effect on poly(A) site choice in many genes. To test this, we first monitored the effect increased levels of CstF-64 in RAW macrophages had on poly(A) site

choice by using a stably transfected reporter construct that contained two linked poly(A) sites. We observed an increase in the use of the weaker promoter-proximal poly(A) site in the reporter construct concomitant with an increase in CstF-64 protein levels following LPS stimulation. Because of the pleiotropic effects LPS stimulation has on macrophages, we stably over-expressed CstF-64 in RAW macrophages to ascertain if CstF-64 specifically influences poly(A) site choice. Indeed, RAW macrophages that stably over-express CstF-64 also demonstrated an increase in weak poly(A) site choice on the reporter construct. We therefore hypothesized that induction of CstF-64 by LPS has a functional consequence for gene expression in activated macrophages. To determine how many genes could be influenced, we performed microarray analysis of gene expression from RAW macrophages stably over-expressing CstF-64. We found that 10-fold constitutive over-expression of CstF-64 in RAW macrophages significantly altered the expression of 52 genes, of which over 25% share common gene expression changes with LPS-stimulated RAW macrophages that exhibited a 5-fold maximal increase in CstF-64 expression. Closer analysis of five of the genes whose expression changed under both LPS-stimulating and CstF-64 over-expressing conditions, Id-2, CTE-1, MMP-9, Tyki and Fabp-4, revealed that an alternative polyadenylation event does occur on these gene transcripts and that this change in poly(A) site choice may contribute to the abundance of mature transcripts by the removal of mRNA instability or addition of mRNA stability elements. From this, we conclude that increases in expression of the pre-mRNA cleavage factor CstF-64 observed in LPS-stimulated macrophages significantly contribute to changes in the expression of a multitude of genes through alternative polyadenylation events that occur in the context of infection by gram-negative bacteria.

4.3. MATERIALS AND METHODS

4.3.1. Cell Culture and Treatment

Murine BMDMs were derived from C57BL/6 bone marrow based on adherence. In brief, the bone marrow from euthanized mice was flushed out of the femur and tibia bones with Dulbecco's modified Eagle medium (DMEM)(Life Tech.) as previously described (13). The bone marrow suspension was then washed twice with 2% FBS (Atlanta Biologicals, Norcross, GA) in PBS. Non-tissue culture-treated petri dishes (Labtek) were seeded at 2×10^6 cells per dish in 25 ml of macrophage media (25% L-cell supernatant, 20% fetal bovine serum, 1% L-glutamine, 1% pyruvate, and 1% nonessential amino acids). After 4 days in culture the cells were fed with 10 ml of fresh macrophage media. BMDMs were grown until approximately 80% confluent, then treated with 100 ng/ml LPS (Sigma, strain O111:B4) for 18 hours.

RAW 264.7 murine macrophages (ATCC) were maintained in DMEM with 10% FBS, 10 mM HEPES, and 0.1% penicillin-streptomycin. Cells were grown in culture flasks until approximately 80% confluent, then treated with 100 ng/ml LPS for 18. Stably transfected polyclonal cell lines were also maintained under G418-neomycin (Cellgro) and/or mycophenolic acid (MPA) (GIBCO-BRL) antibiotic pressure, depending on the incorporated plasmid(s).

4.3.2. Western Blot Analysis and Antibodies

Ten micrograms of whole cell lysates were loaded per lane unless otherwise indicated. Proteins were visualized after treatment of the appropriate secondary antibody with the Renaissance™ Chemiluminescence system (NEN) according to manufacture's instructions. All Western blots were re-probed with GAPDH to demonstrate equal loading between lanes. Quantification of signal intensities was performed on the Kodak Imaging Station 2000R and associated 1D analysis software. CstF-64 expression was normalized to the GAPDH signal on

the same blot. Rabbit antiserum recognizing murine CstF-64 was generated for peptide CIAMLPPEQRQSILILKEQIQKSTGAP, corresponding to the 26 C-terminal amino acids (a cysteine residue was added to the N-terminus to aid KLH coupling). Injections and collections of serum were performed at Charles River Pharmservices. The anti-CstF-50 and anti-CstF-77 rabbit anti-peptide antibodies were made in our laboratory as previously described (86) as well as anti-hnRNP F and anti-hnRNP H/H' rabbit polyclonal antibodies (146). Anti-CPSF-30 (rabbit polyclonal) and anti-CPSF-100 (rabbit polyclonal) were generously provided by Walter Keller (University of Basel). We purchased the mouse monoclonal anti-glyceraldehyde-3-phosphate dehydrogenase (GAPDH) antibody from Chemicon. HRP-conjugated secondary antibodies for mouse IgG (Sigma A7282) and rabbit IgG (Sigma A0545) were used according to manufacturer's instructions.

4.3.3. Measurement of Cell Proliferation by [^{3}H] Thymidine Incorporation

RAW macrophages were plated in 35-mm culture dishes and allowed to grow to ~80% confluency (3 x10 ^{6} cells/plate) with 5 ml of medium and left untreated or stimulated with LPS. To measure DNA synthesis (99), 10 μCi of [^{3}H]dT (6.7 Ci/mmole) in aqueous solution was added to the cells covered with 3 ml of medium. After 6, 14.5 and 18 hours at 37^0 C, the medium was removed and the cells were washed twice with 1X PBS and then trypsinized. Cells were collected, spun down and resuspended in 1 ml PBS; cell number was determined from an aliquot. Triplicate 200-μl samples were removed. Nucleic acids were precipitated by 10% trichloroacetic acid (TCA) and collected on glass fiber filters. Radioactivity was determined using scintillation counting and the [^{3}H]dT incorporated was normalized for cell counts.

4.3.4. RNA Isolation and Poly(A)$^+$ RNA Purification

Total RNA was extracted from RAW macrophages using the Ultraspec RNA isolation system (Biotecx, Houston, TX) according to the manufacturer's instructions. Poly(A)$^+$ RNA was purified from total RNA using Oligo(dT) Cellulose (Stratagene) according to manufacturer's instructions. Purified mRNA was suspended in DEPC (Sigma) treated water and 2 volumes 95% ethanol and stored at -20°C until needed.

4.3.5. T2 RNase Protection Assay

RAW macrophages were stably transfected with a guanine phosphotransferase (gpt)-containing plasmid with linked poly(A) sites and cultured in the presence of MPA, as previously described (87). The gpt gene has a weak poly(A) site from the α2-globin gene upstream of the strong SV40 late poly(A) site that is contained in the pSV2gpt vector (95). The RAW macrophages that contain the gpt-encoding reporter gene are subsequently referred to as RAW-α2 macrophages. RNase protection assays were performed using an anti-sense radiolabeled riboprobe that hybridizes to the 3'-UTR of the gpt mRNA to detect poly(A) site usage (see Figure 5A). The riboprobe was synthesized with a specific activity of ~7.3x10^7 cpm/μg as previously described (87). Radioactivity contained in the proximal and distal poly(A) site protected fragments was measured with a PhosphorImager (MD 860, Molecular Dynamics, CA) and quantification was performed using ImageQuant software (Molecular Dynamics). The proximal/distal ratio (P/D) was computed using Office XP Excel (Microsoft) and all ratios from experimental samples were normalized to the untreated control within each experiment. Statistical significance of the data was calculated using SEM and the paired student's t-test on from two to three trials performed from at least two separate RNA preparations, for a total of five trials from each condition.

4.3.6. Over-expression of CstF-64 in RAW 264.7 Murine Macrophages

The complete open reading frame of human CstF-64 [CstF-2] was cloned into the EcoR I sites of pBluescript II SK (+) (pBSSK-64) (CAM 375) using EcoR I linkers as described (134) and was a gift from Jim Manley (Columbia University). Having difficulties with CMV promoter activity in RAW 264.7 macrophages, we used pEF1/Myc-His B (Invitrogen), a plasmid driven by the human EF-1α promoter that has been shown to achieve high levels of exogenous gene expression in RAW macrophages (72). A double digest of both pBSSK-64 and pEF1/Myc-His B with Kpn I (Boehringer Mannheim) and Xba I (Boehringer Mannheim) was performed followed by gel purification of an ~2.0 kb fragment from pBSSK-64 and an ~6.5 kb fragment from pEF1/Myc-His B. Ligation of the purified plasmid fragments with T4 DNA ligase (NEB) was done in a 16°C waterbath overnight followed by direct transformation of the ligation reactions into One Shot MAX Efficiency DH5α-1 competent cells (Invitrogen) resulting in a sufficient number of colonies on Penicillin G (Sigma) LB plates for miniprep screenings. The resulting CstF-64 expression plasmid is called pEF1-64 (CAM 508).

RAW-α2 macrophages were stably transfected using FuGene 6 (Roche) with pEF1-64 or pEF1-lacZ, as a control, as instructed by the manufacturer. In brief, RAW-α2 macrophages were seeded in 6-well culture dishes and grown until approximately 75% confluent. Macrophages were transfected with a FuGene 6 (μl):DNA ratio (μg) of 8:1 in the presence of serum. At day 3 post-transfection, the cells were split and placed under selective pressure of 400 μg/ml G418 sulfate (Cellgro) and MPA. Media was changed as needed. After 18 days of selective pressure, poly-clonal populations of RAW-α2-64 or RAW-α2-LacZ cells were stored in liquid nitrogen. Raw-α2-64 cultures were assessed for constitutive CstF-64 expression by Western blot.

4.3.7. Affymetrix Gene Chip Analysis

Three unique samples of RNA from each treatment (untreated, 100 ng/ml LPS treated for 18 hours and constitutively over-expressed CstF-64) were used to create biotinylated cRNA at the PittArray DNA microarray facility of the Genomics and Proteomics Core Laboratories (GPLC) of the University of Pittsburgh. Fifteen micrograms of fragmented RNA were hybridized to the murine 430A Affymetrix gene chip according to manufacturer's instructions. Affymetrix gene chips were scanned using MAS 5.0 (Affymetrix) to obtain raw expression values and signal calls (i.e. Present, Absent or Marginal) for each probe set on the array chip (GEO Accession No. GSE2002). Data sets of intensities for 22,690 probe sets per expression array were analyzed by BRB-ArrayTools v3.02 (http://linus.nci.nih.gov/BRB-ArrayTools.html). Data was filtered through the software using the following parameters: all signals with an intensity value below 10 were given a threshold value of 10 and each array was normalized to the 64 Affymetrix control probe sets intrinsic to the murine 430A chip. Furthermore, probe sets were excluded from all arrays under the following conditions: when less than 30% (i.e. 0, 1, or 2 of 9) of expression data values have a 2-fold change in either direction from the probe set's median value or when 50% or more of the data for a probe set is missing or absent. Using these filtering parameters, 1145 probe sets passed the criteria. To identify genes that demonstrated statistically significant changes in expression between control and treatment (LPS or CstF-64 over-expressing) groups, the probe sets that had passed the initial filtering process also had to pass the following criteria: 1) the mean of the treatment group must have demonstrated a 2-fold increase or decrease in signal over the mean of the control group, 2) the signal values for each probe set must have passed a statistical univariant significance test with p-value<0.05 with 1000 permutations and specifying a random variance for each calculation (Class Comparison feature of BRB Array Tools), and 3) a probe set was excluded if it scored an Absent call in at least 2 out

of 3 measurements for the treatment group (for up-regulated genes) or scored an Absent call in at least 2 out of 3 measurements for the control group (for down-regulated genes). Marginal calls were considered Absent for this criterion.

A statistical algorithm was designed to test if the genetic overlap between LPS-stimulated and CstF-64 over-expressing conditions was beyond the random chance of occurrence. Briefly, the test was designed with one set consisting of a draw of 515 random numbers from a pool of 22,690 and another set consisting of a draw of 52 random numbers again from a pool of 22,690 numbers. The test was run 30,000 times looking for the number of random chance over-laps from the two sets of randomly selected numbers. A graph depicting the frequency of zero numbers of overlaps, one number, two numbers, etc. up to fourteen numbers from 30,000 trials was constructed using Microsoft Excel and appears in Figure 10.

4.3.8. Semi-quantitative RT-PCR

Single-stranded cDNA was synthesized for each condition from 10 μg total cellular RNA using the Superscript First-Strand Synthesis System for RT-PCR with oligo dT primers (Invitrogen). PCR reaction mixtures were assembled in quadruplicate for each primer pair using either 0.5 or 1.0 μl cDNA, 50 pmol of each primer, 800 nM dNTP mix, 5 μl 10X Taq buffer and 2.5 Units Taq (Genechoice) in 50 μl volumes. At the conclusion of 20, 25 and 30 cycles, reaction tubes were removed from the PCR machine and placed on a heat block for a 7 minute, 72°C final extension. The last tube of each set of four reactions was allowed to complete the PCR program to 35 cycles on the thermocycler. Ten microliter aliquots of the PCR reactions representing 20, 25, 30 and 35 cycles were loaded on 1.5% agarose gels, stained with ethidium bromide, and visualized using the Kodak Imaging Station 2000R. The HPRT gene, which is shown to be equally expressed between the RAW macrophage culture conditions on the

microarray data (GSE2002) and has been previously shown to be a good candidate for RT-PCR loading control (49) was used to demonstrate equal cDNA concentrations between preparations. The nucleotide sequences for the primer pairs used for semi-quantitative RT-PCR are listed in Table 1. Primer sequences were designed from GenBank cDNA sequences using Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) unless otherwise indicated.

4.3.9. PCR-based Identification of Alternative Polyadenylation (MRP-PCR)

A multiple reverse primer PCR (MRP-PCR) approach was implemented to identify a poly(A) site switching event in the 3'-UTR of selected genes. This MRP-PCR approach used one primer pair that lies upstream of all putative poly(A) sites that would amplify all gene transcripts regardless of the poly(A) site usage. A second anti-sense primer that targets the downstream sequence between two putative poly(A) sites was used in conjunction with the common sense primer to identify the abundance of gene transcript that had been processed at the distal poly(A) site. The longest GenBank sequence for each gene that demonstrated significant 2-fold or greater changes under CstF-64 over-expressing conditions was scrutinized for possible alternative poly(A) sites and AU-rich elements. A putative poly(A) site was considered one that contains either a consensus AAUAAA or non-consensus AUUAAA poly(A) signal followed 30-60 nucleotides downstream by a GU-rich element that possess at least two contiguous uridines. Possible AU-rich regulatory elements were based on the functional AU-rich motifs listed in the review by Carol Wilusz *et al*. (150). Using these criteria, five candidate genes were chosen to test for the possibility of alternative polyadenylation within their 3'-UTR: Id-2, CTE-1, MMP-9, Tyki, and Fabp-4. Single stranded cDNA from each condition was subject to PCR using a sense primer (F) and an antisense primer (R1) that lies upstream to all putative poly(A) sites or an

antisense primer (R2) that lies between the putative proximal and putative distal poly(A) sites (Figure 11). The nucleotide sequences used for MRP-PCR are listed in Table 4.

4.4. RESULTS

4.4.1. LPS stimulation of RAW macrophages specifically increases CstF-64 protein expression.

We treated RAW macrophages with increasing amounts of LPS for 18 hours and measured CstF-64 protein expression by Western blot. We observed that CstF-64 protein levels increased with LPS stimulation of RAW macrophages and that this increase is LPS dose-dependent (Figure 3A). The induced expression of CstF-64 plateaued at 100 ng/ml LPS (Figure 3A). A semi-quantitative Western blot revealed that 100 ng/ml LPS stimulation of RAW macrophages for 18 hours resulted in a 4-5 fold increase in CstF-64 expression (Figure 3B). Quantification of the luminescence from the Western blot using the Kodak Imaging Station 2000R corroborated the semi-quantitative Western data by measuring the induction of CstF-64 with 100 ng/ml LPS at 5-fold that of unstimulated RAW 264.7 macrophages (see Methods). Therefore, we conclude that LPS treatment of RAW macrophages increased CstF-64 protein expression approximately 5-fold. From these observations, 100 ng/ml LPS treatment for 18 hours was used for all subsequent experiments.

LPS stimulation of B lymphocytes not only increases CstF-64 expression (139), but also induces cell cycle progression resulting in rapid clonal expansion (117). We have previously shown that CstF-64 protein levels increase when serum-starved, resting mouse 3T3 fibroblasts were induced into the S-phase of the cell cycle by the addition of serum-containing media (86). Both of these examples demonstrate that CstF-64 protein levels increase in the presence of cell cycle progression. Notably, LPS stimulation of macrophages induces cell cycle arrest (144, 145). By monitoring the incorporation of [^{3}H] thymidine under LPS stimulating conditions, we confirmed cell-cycle arrest upon LPS treatment of RAW macrophages (Figure 3C). Therefore,

we conclude that LPS stimulation of RAW macrophages induces a 5-fold increase in CstF-64 protein expression that is not obligatorily coupled to the cell cycle.

The protein levels of several other cleavage/polyadenylation factors were also assessed under LPS stimulating conditions. Western blots showed no measurable change for CPSF-30, CstF-77, and CF-I_m-68 in LPS stimulated RAW macrophages (Figure 4, left panel). Furthermore, the expression of hnRNP H/H' and hnRNP F, RNA processing factors we have found to have effects on cleavage/polyadenylation in B lymphocytes and plasma cells (146), also remained unchanged (Figure 4, left panel). To determine if the specific increase in CstF-64 protein expression observed in the transformed RAW cell line also occurred in primary macrophages, we stimulated murine bone marrow derived macrophages (BMDMs) with 100 ng/ml LPS for 18 hours. CstF-64 protein was increased approximately 10-fold in murine BMDMs stimulated with 100 ng/ml LPS (Figure 4, right panel). As with the RAW macrophage cell line, no other mRNA processing factors assayed changed expression upon LPS stimulation (Figure 4, right panel). Studies have shown that CstF-64 is the limiting component for CstF trimer formation, the active form of CstF (86, 139). Therefore, we hypothesized that the specific increase in CstF-64 observed in LPS stimulated RAW macrophages and BMDMs has a significant impact on 3'-mRNA processing and would thus alter the expression of specific genes.

4.4.2. LPS stimulation of RAW macrophages induces a poly(A) site switch on a reporter mini-gene.

Previous studies have shown that increases in CstF-64 expression in LPS-stimulated splenic B-lymphocytes are associated with an increase in weak poly(A) site selection on the IgM heavy chain gene (139). With the specific increase in CstF-64 protein observed in LPS-stimulated macrophages (Figure 4), we hypothesized that the nuclear mRNA processing

machinery has a heightened level of mRNA 3'-end processing in activated macrophages that results in an increased selection of weak poly(A) sites. To test this, we stably transfected into the RAW macrophage cell line a reporter mini-gene that possesses two contiguous poly(A) sites of differing strengths in its 3'-UTR (Figure 5A). The reporter mini-gene codes for the bacterial gene guanosyl phosphotransferase (gpt) and is under the control of the SV40 early promoter (95). Stable incorporation of the mini-gene in the RAW macrophages is thus conferred through mycophenolic acid (MPA) resistance. Contained in the 3'-UTR of the mini-gene is the weak α2-globin poly(A) site upstream of the strong SV40 early poly(A) site (Figure 5A). The RAW macrophage cell line that stably expresses the α2-globin poly(A) site containing gpt min-gene is referred to as RAW-α2. LPS-stimulation of RAW-α2 macrophages exhibited the same protein expression levels for CstF-64 and the other cleavage/polyadenylation factors measured in untransfected RAW macrophages (Figure 4 and data not shown). Previous studies using this same reporter mini-gene in different B-cell lines demonstrated that the upstream, weaker poly(A) site was selected more often in the plasma cell environment which exhibits heightened 3'-end processing activity than in memory B-cell lines (87). Here, we wished to test the hypothesis that the LPS-stimulation of RAW macrophages that exhibit elevated levels of CstF-64 expression induces an increase in weak promoter-proximal poly(A) site selection on the reporter mini-gene.

We performed T2 RNase protection assays on mRNA isolated from unstimulated or LPS-stimulated RAW-α2 macrophages to measure poly(A) site selection within the 3'-UTR of the gpt reporter mini-gene, as previously described (87). The full-length probe is 564 nucleotides in length. A protected fragment 487 nucleotides in length indicates that the promoter-distal poly(A) site was chosen for mRNA cleavage and polyadenylation (Figure 5A). Likewise, a protected fragment 185 nucleotides in length indicates the choice of the promoter-proximal poly(A) site for

mRNA processing. Importantly, the poly(A)-containing RNA for the T2 RNase protection assay was harvested at 18 hours after LPS stimulation. This allowed the cells to achieve the 5-fold increase in CstF-64 protein expression we had previously observed (Figure 3B). A representative result of the RNase protection experiment displays the protected, radio-labeled riboprobe fragments separated on a denaturing urea gel (Figure 5B). Quantification of the amount of transcripts corresponding to each protected fragment was performed using a Phosphorimager and the ImageQuant software package (see Methods). By calculating the proximal to distal ratio (P/D) of the protected fragments, we were able to obtain a relative measurement of poly(A) site usage that can be compared between experiments and treatments. An increase in P/D for a particular treatment indicates that there is an increase in selection of the weaker promoter-proximal poly(A) site of the gpt mini-gene transcripts relative to the stronger promoter-distal poly(A) site. Likewise, a decrease in P/D would indicate increased selection of the stronger promoter-distal poly(A) site. Although the absolute values of transcripts processed at the upstream and downstream poly(A) sites vary between experiments, the P/D ratio between samples in a particular experiment were reproducible. By normalizing the P/D for LPS-treated RAW-α2 macrophages relative to the unstimulated control, we measured a 2.4-fold increase in the P/D in macrophages stimulated with LPS indicating an increase in upstream, weak poly(A) site selection (Figure 5C). Because of the steady-state pool of mature gpt transcripts that is in the cell prior to LPS-induced elevation of CstF-64 protein, we believe that the observed 2.4-fold increase in P/D is a conservative measurement of increased mRNA 3'-end processing associated with LPS-induced CstF-64 protein levels. Elevated levels of CstF-64 therefore correlate with an increased selection of the promoter-proximal poly(A) site in LPS-treated RAW-α2 macrophages.

4.4.3. Over-expression of CstF-64 in RAW macrophages induces a poly(A) site switch on a reporter mini-gene.

Over-expression of CstF-64 in chicken B-lymphoma DT40 cells demonstrated an increase in weak poly(A) site selection of the Ig heavy-chain gene, resulting in an increase in the secretory form of the IgM transcript (139). Likewise, over-expression of CstF-64 in NIH-3T3 cells increases promoter-proximal poly(A) site selection in Testis Brain RNA-binding protein (24). In order for us to be certain that the observed 5-fold increase in CstF-64 protein expression observed in LPS-stimulated RAW-α2 macrophages is responsible for the switch in poly(A) site selection on the gpt transcript, we stably over-expressed CstF-64 in RAW-α2 macrophages. Using a mammalian expression vector driven by the human EF1α promoter, we achieved a stable 10-fold increase in CstF-64 protein expression in RAW-α2 macrophages as quantified using the Kodak Imaging Station 2000R (see Methods and Figure 6A). Cells exhibiting constitutive over-expression of CstF-64 in RAW-α2 macrophages are referred to as the RAW-α2-64 macrophage cell line. Because it was the human form of CstF-64 that was constitutively over-expressed in the murine RAW 264.7 macrophage cell line, it is important to explain the differences between the mouse and human forms of CstF-64. There is 94% amino acid sequence homology between the two forms of the protein (Figure 7). More importantly, the functional portions of the protein have a much greater degree of homology. For example, the RRM of the human protein is an exact match to the mouse form. Furthermore, the hinge region that is responsible for binding to CstF-77 (136) has only one amino acid difference (serine to threonine; S118T) as does the carboxy-terminal domain of the protein (asparagine to serine; N528S) that as been shown to bind PC4 (18). One other notable difference between the human and mouse form of CstF-64 is a triple proline insert that appears in the mouse form and is located downstream of

Proline 350. The effect of this extra proline triplet in the mouse protein has not been functionally characterized. However, because the proline triplet lies within the less conserved proline/glycine-rich region that has yet to be functionally characterized, we believe the lack of these prolines in the over-expressed human form of CstF-64 will not detract from our studies of cleavage and polyadenylation in the murine cell line. Because of the sequence similarities between the human and mouse forms of CstF-64, we believed that using the human form of CstF-64 in the murine RAW 264.7 macrophage cell line was an acceptable approach to achieve CstF-64 over-expression.

Western blots for the same array of proteins measured in the LPS studies (Figure 4) revealed no significant changes in these proteins under CstF-64 over-expressing conditions (data not shown). In the RAW-α2-64 macrophages, we measured the effect increased levels of CstF-64 protein had on poly(A) site selection by T2 RNase protection assays, as previously performed in the LPS-stimulation studies. From our observations that LPS stimulation of RAW-α2 macrophages increased CstF-64 protein expression and induced an increase in weak promoter-proximal poly(A) site selection of our gpt reporter mini-gene, we hypothesized that over-expression of CstF-64 alone would induce some observable increase in weak poly(A) site selection of the gpt reporter mini-gene. We found that in the RAW-α2 macrophages that constitutively over-expression CstF-64, there was a 2.0-fold increase in the P/D ratio compared to unstimulated RAW-α2 macrophages (Figure 6B). From these observations, it is clear that increased levels of CstF-64 protein alone have as significant an impact on poly(A) site choice as that measured under LPS stimulating conditions.

We conclude that elevated levels of CstF-64 protein expression in LPS-stimulated macrophages contributes to an increase in weak poly(A) site selection in RAW macrophages on

the gpt reporter mini-gene. From this, we reasoned that elevated levels of CstF-64 might have an effect on macrophage gene expression in the context of LPS stimulation.

4.4.4. Microarray analysis reveals that forced over-expression of CstF-64 in RAW macrophages changes the expression of multiple genes.

We have demonstrated that both LPS stimulation and CstF-64 over-expression induced an increase in the selection of weak poly(A) sites in RAW macrophages on a reporter mini-gene that possesses two contiguous poly(A) sites of varying strengths. Because of the impact that post-transcriptional processing has on gene expression, we compared global gene expression profiles in unstimulated, LPS-stimulated, and CstF-64 over-expressing RAW macrophages using the microarray technique. RNA from untreated and CstF-64 over-expressing RAW macrophages and RAW macrophages stimulated with 100 ng/ml LPS for 18 hours was used for microarray analysis. Using Affymetrix murine gene chip 430A, 22690 genes were examined from 3 replicates of each treatment group (see Methods). Statistical analysis was performed as described in Methods using BRB Array Tools developed by the National Cancer Institute (http://linus.nci.nih.gov/BRB-ArrayTools.html).

Treatment of RAW macrophages with 100 ng/ml LPS for 18 hours resulted in a statistically significant 2-fold or greater increase in the expression of 245 known genes and 45 expressed sequence tags (ESTs) and a 2-fold or greater decrease in the expression of 162 known genes and 63 ESTs (Appendix A and B). To ensure that our data sets from LPS-stimulated macrophages accurately represent an activated macrophage gene expression profile, we performed semi-quantitative RT-PCR using the same RNA sample used for the microarray experiment to measure TNF-α,.iNOS and COX-2 transcript levels, all of which are identifiers of macrophage activation. As with the microarray data, we have shown that TNF-α, iNOS and

COX-2 transcript levels all significantly increased in the presence of LPS stimulation (Figure 8). We also confirmed the change in expression of several other genes by RT-PCR (Figure 8). None of the CPSF subunits, poly(A) polymerase (PAP), or poly(A) tract binding protein (PABP) changed expression upon LPS stimulation (Appendix A and B). Probes for CstF-77, CstF-50, hnRNP F, hnRNP H/H' or CF-Im-68 were not included on the Affymetrix 430A murine gene chip.

Microarray analysis of RAW-α2 macrophages over-expressing CstF-64 protein approximately 10-fold over basal levels showed a statistically significant 2-fold or greater increase in the expression of 21 known genes and 2 ESTs, of which 12 of the genes demonstrated a significant increase in LPS stimulation (Table 2 and Figure 9). Likewise, 29 known genes showed a 2-fold or greater decrease in the expression with CstF-64 over-expression, of which 3 of the genes also demonstrated a significant decrease with LPS stimulation (Table 3 and Figure 9). We developed a statistical algorithm that tested the chance that the selection of 15 genes from each experimental pool (i.e. LPS-stimulated and CstF-64 over-expressed) would occur at random (see Methods). Our results clearly show that the 15 genes that change expression in both LPS-stimulated and CstF-64 over-expressing RAW macrophages did not occur at random and was thus a real biological observation (Figure 10). Analysis of the 3'-UTRs of all 52 genes that change expression under CstF-64 over-expression revealed that over 46% of them (24 of 52) possess two or more putative poly(A) sites (Table 2 and 3). Putative poly(A) sites were chosen by their consensus AAUAAA or nonconsensus AUUAAA poly(A) signal (shown to be greater than 58% and 14%, respectively, of the functional poly(A) signals determined from human database analysis (9)) and an associated GU-rich region that possesses at least two contiguous uridines which lies within 20-60 nucleotides downstream of the poly(A) signal. Many of these

genes that may support an alternative polyadenylation event also possess adenosine/uridine (AU)-rich elements (AREs) that may, themselves, dictate mRNA stability or translatability [reviewed in (150) and (157)].

Notably, we did not detect increases in CstF-64 gene expression in LPS-stimulated RAW macrophages by microarray analysis. This result was not unexpected because repeated attempts to measure changes in CstF-64 gene expression in LPS-stimulated RAW macrophages by Northern blot or RT-PCR demonstrated no change in CstF-64 message (data not shown). This result led us to believe that the observed increase in CstF-64 protein is the result of a post-transcriptional/translational control mechanism, a characteristic of other genes in LPS-stimulated macrophages (27, 157). Furthermore, we were unable to measure increases in CstF-64 gene expression in RAW-α2-64 macrophages by microarray because 9 out of the 11 probe pairs in the CstF-64 probe set for the murine 430A gene chip target 3'-UTR genomic sequences (Affymetrix NETAFFX™ Analysis Center, http://www.affymetrix.com/analysis/index.affx) which are not contained in the CstF-64 expression vector that is stably incorporated into the RAW macrophage genome. Nevertheless, the Western blot data for both LPS-stimulated (Figures 3A and 3B) and CstF-64 over-expressing (Figure 6A) RAW macrophages firmly established that protein levels are indeed increased over basal levels under both conditions.

Therefore, we conclude that increased expression of CstF-64 contributes, in part, to the gene expression changes that accompany the late macrophage response to LPS stimulation.

4.4.5. Increased levels of CstF-64 induce alternative polyadenylation on several LPS-responsive genes in RAW macrophages

Interested in whether elevated levels of CstF-64 in the RAW macrophage influence the poly(A) site choice of LPS-responsive genes, we set out to detect by PCR a possible poly(A) site

switch on genes that possess multiple putative poly(A) sites, have an AU-rich element (ARE), and whose expression changes in both LPS-stimulated and CstF-64 over-expressing RAW macrophages. Figure 11A describes our multiple reverse primer PCR (MRP-PCR) approach to detect the use of alternative poly(A) sites under LPS-stimulating and CstF-64 over-expressing conditions. By pairing the common forward primer (F) with the upstream reverse primer (R1), we were able to detect all of the mature transcripts regardless of which poly(A) site was chosen. Furthermore, by pairing the F primer with the downstream reverse primer (R2), we were able to detect the portion of the mature gene transcript that was cleaved and polyadenylated at a more distal poly(A) site (FR2).

Interestingly, for Id-2, CTE-1 and MMP-9, we were able to detect a reduction in FR2 PCR product compared to FR1 PCR product, which represents total gene expression, under LPS-stimulating conditions compared to the untreated control (Figure 11B). Likewise, we also detected a reduction in FR2 PCR product for Id-2 and MMP-9 under CstF-64 over-expressing conditions compared to FR1 PCR product (Figure 11B). Since we demonstrated an increase in the total gene transcript using the FR1 primer pair for Id-2 and MMP-9 under LPS-stimulating and CstF-64 over-expressing conditions, this proves that the decrease in FR2 PCR product observed is the result of an increase in proximal poly(A) site selection, an event similar to what was observed on our reporter mini-gene (Figure 5). Because total CTE-1 expression is greatly reduced under CstF-64 over-expressing conditions (Table 3 and Figure 8), we were unable to determine if loss of the FR2 PCR product was due to an authentic poly(A) site switch or that the switch did not occur and the FR2 PCR product was merely undetectable due to low abundance of over-all gene product.

Unexpectedly, we observed an increase in FR2 PCR product under LPS-stimulating conditions for Fabp-4 and Tyki (Figure 11C). Likewise, a similar increase in FR2 PCR product was observed under CstF-64 over-expressing conditions (Figure 11C). Counterintuitive to our observation of increased proximal poly(A) site selection under elevated levels of CstF-64 (Figures 5 and 11B), these results clearly demonstrate that there is an increase in the distal poly(A) site selection for Fabp-4 and Tyki under LPS-stimulating conditions and this can be contributed, in part, to elevated levels of CstF-64.

4.4.6. Genes that demonstrate alternative polyadenylation with CstF-64 increases possess putative regulatory elements within their 3'-UTR.

By identifying a poly(A) site switch in both Id-2, CTE-1, MMP-9, Tyki and Fabp-4 under LPS-stimulating conditions, we focused our attention on the region between the proximal and distal poly(A) sites of these gene transcripts in the hopes to gain some insight on a biological effect a change in poly(A) site use would have on these genes' expression. In Figure 12, we highlight the putative AREs of all five genes that lies between the proximal and distal poly(A) sites. Interestingly, all of these regions possess AU-rich elements and/or U-rich regions that resemble biologically functional AREs that have been previously described to influence mRNA half-life and/or translation [(150) and Table 5].

From this, we conclude that the physiological increase in CstF-64 protein through LPS stimulation causes, in part, a shift in poly(A) site usage on LPS-inducible genes and that this shift in poly(A) site usage may contribute to the final amount of the gene transcript and influence the translatability of that transcript, thus profoundly effecting the amount of protein produced.

4.4.7. Post-transcriptional regulation of CstF-64 in LPS-stimulated RAW macrophages

We reported an increase in CstF-64 protein approximately 5-fold over basal levels in RAW 264.7 murine macrophages stimulated with LPS (Figure 3). However, we did not detect an increase in CstF-64 gene transcript with LPS stimulation by Northern Blot, RT-PCR (data not shown), or microarray (Appendix A). Because we have consistently observed increases in protein expression in the absence of mRNA increases, we hypothesize that the increase in CstF-64 protein results from either the accumulation of CstF-64 stemming from an increase in the protein half-life or an increase in translation of the steady-state mRNA under LPS-stimulating conditions. Closer scrutiny of the 3'-UTR of CstF-64 mRNA revealed that it contains two poly(A) sites with an ARE core sequence and associated U-rich elements that lie immediately upstream of the distal poly(A) site (Figure 13). Interestingly, previous work from our lab has shown by Northern Blot that there are two main species of CstF-64 mRNA in mouse B-lymphocytes and that the increase in a short 2.2 kb form is observed over a longer 4.2 kb form in plasma cells relative to naïve and memory B-cells (Souan MS thesis). Support for the existence of two forms of CstF-64 mRNA also stems from two GenBank entries for the mouse form of CstF-64 that exhibit different lengths. The shorter form of the mouse CstF-64 transcript (Accession AF317552) uses a poly(A) site that creates an approximate 2.1 kb form of the transcript. The longer form of the mouse CstF-64 transcript (Accession NM_133196) uses a distal poly(A) that creates an approximate 4.4 kb transcript. Because we have not assessed the abundance of CstF-64 mRNA by Northern Blot in RAW 264.7 macrophages, we cannot say whether there exist alternative forms of the message that stem from a change in poly(A) site usage occurring under LPS-stimulating conditions. However, because the amount of mRNA transcript is unchanged under LPS-stimulating conditions, we hypothesize that there are *cis*-elements within the transcript that control translation which is responsible for the increased

protein production observed with LPS stimulation. Northern Blot analysis of CstF-64 message in RAW macrophages would provide evidence for multiple CstF-64 transcripts and MRP-PCR would determine if a poly(A) site switch occurs under LPS-stimulating conditions.

4.5. DISCUSSION

We have shown that LPS stimulation of murine RAW macrophages stimulated with LPS for 18 hours exhibited CstF-64 expression levels approximately 5-fold over basal expression, and occurred during the cessation of the cell cycle (Figure 3**)**. This finding is noteworthy due to studies from our lab and others that have shown that CstF-64 expression levels increase with cell cycle progression (86, 139). Furthermore, we showed that LPS stimulation of both RAW macrophages and murine BMDMs increased CstF-64 specifically, in that several other proteins associated with pre-mRNA cleavage/polyadenylation did not change expression (Figure 4).

Through RNase protection assays, we measured a 2.4-fold increase in the selection of the weaker promoter-proximal poly(A) site of the reporter mini-gene in LPS-stimulated RAW macrophages that exhibit increases in CstF-64 protein expression (Figure 5). Constitutive over-expression of CstF-64 approximately 10-fold higher than basal levels further demonstrated that specific increases in CstF-64 protein expression were sufficient to induce a 2.0-fold increase in selection of the weaker promoter-proximal poly(A) site on the reporter mini-gene (Figure 6).

Microarray analysis revealed that LPS stimulation of RAW macrophages for 18 hours significantly increased the expression of 245 known genes and 45 unique ESTs (Appendix A). We also identified a significant decrease in the expression of 162 known genes and 63 unique ESTs in LPS-stimulated RAW macrophages (Appendix B). This data demonstrates that the effects of LPS stimulation on macrophage gene expression are sustained beyond the initial induction of gene transcription and thus could be contributed to by a variety of secondary effects. Furthermore, we measured a significant increase in 21 known genes and 2 ESTs, and a significant decrease in 29 known genes in RAW macrophages that over-express CstF-64. Because the only identified function of CstF-64 to date is its required role in the 3'-end cleavage

step of pre-mRNA processing, we propose that genes that exhibit expression increases in the context of CstF-64 over-expression do so because: (1) more gene transcript is cleaved and polyadenylated from the nascent pool of pre-mRNA of that gene, especially in the context of a single, weak poly(A) site; (2) an alternative polyadenylation event occurs that removes mRNA instability elements from the mature mRNA transcript; or (3) there are secondary effects from genes that regulate transcription which are themselves directly effected by elevated levels of CstF-64 (Figure 9). Notably, there are at least three members in our list of CstF-64 inducible genes that have known transcriptional regulatory functions: CHOP/GADD153, Id2 and retinoic acid receptor beta (RARβ). CHOP/GADD153 is a member of the C/EBP transcription factor family whose heterodimerization with other C/EBP transcription factors results in dominant negative inhibition. However, contact with AP-1 containing complexes induces AP-1 associated gene expression (121, 142). Id2 is a member of the bHLH family of transcription factors that lacks a DNA binding motif and represses transcription when heterodimerized to other bHLH-containing transcription factor. Id2 has also been shown to be upregulated in macrophages under hypoxia-induced stress (16, 133). Retinoic acid receptor beta (RARβ) has been found to be a constitutive activator of retinoic acid response element (RARE)-containing genes irregardless of the presence of its ligand retinoic acid (62). Likewise, we believe the genes that exhibit a decrease in expression with CstF-64 over-expression are influenced by: (2) an alternative polyadenylation event that removes an mRNA stability element from the mature mRNA transcript, or (3) the transcriptional regulators that increase in expression under elevated levels of CstF-64 repress the expression of these genes (Figure 9).

Interestingly, 12 of the 23 genes that increased and 3 of the 29 genes that decreased in the context of CstF-64 over-expression also demonstrated similar gene expression changes with LPS

stimulation (Tables 2 and 3). By designing a statistical algorithm that tests the random chance of the same 15 genes out of the 22,690 on the array being selected from two separate pools at random, we found that this overlap between the LPS stimulation and CstF-64 over-expression would occur far less than once in 30,000 trials (Figure 10). We therefore believe that the overlap is highly unlikely to have occurred at random and thus has biological significance. From this we conclude that some of the gene expression changes observed with LPS stimulation of RAW macrophages can be attributed to, at least in part, elevated pre-mRNA cleavage induced by increases in CstF-64 protein expression. A possible explanation of why not every gene that demonstrated expression changes in RAW macrophages over-expressing CstF-64 was not identified in LPS-stimulated RAW macrophages is that the stable 10-fold over-expression of CstF-64 is able to influence the expression of genes that are unaffected by the LPS-induced 5-fold increase. Examples of genes that changed expression with CstF-64 and not with LPS stimulation are PLOD-2, Ph4a2, RARβ, and CD48 (Table 2 and Figure 8). Furthermore, stable over-expression of CstF-64 indefinitely elevates the pre-mRNA cleavage of the cell. This sustained elevation of mRNA 3'-end cleavage acts on all newly synthesized RNA. In the context of LPS stimulation, CstF-64 levels increase to 5-fold over basal expression 18 hours post treatment and therefore heightened levels of pre-mRNA cleavage due to elevated levels of CstF-64 primarily act on gene transcripts produced late in macrophage activation. Transcripts produced before increases in CstF-64 expression, especially those with lengthy mRNA half lives, can dampen the measurable effect elevated levels of CstF-64 have on newly transcribed genes.

Our studies have shown that the mechanism of alternative polyadenylation is strongly influenced by the physiological increase in the expression of CstF-64, resulting in an increase in the selection of a weaker promoter-proximal poly(A) site over a stronger promoter-distal poly(A)

site on a reporter mini-gene. Analyses of EST databases have estimated that as many as 54% of human genes and 32% of mouse genes possess more than one poly(A) site, often of varying strengths (9, 140). Indeed, a multitude of studies have demonstrated how the inclusion or exclusion of regions of the 3'-UTR to the final gene product through alternative polyadenylation can have a profound effect on mRNA half-life (1, 40, 78, 101), the translational efficacy of the gene transcript (88, 157), and the final structure of the protein [(25) and reviewed in (45)]. Analysis of the longest 3'-UTRs available through the GenBank database of the genes that demonstrate changes in abundance with CstF-64 over-expression revealed that 24 out of 52 of these genes possess more than one putative poly(A) site. Furthermore, many of these genes also possess ARE-containing *cis*-elements that may regulate mRNA stability. Therefore, increases in CstF-64 protein levels can, in theory, dictate the presence or absence of these mRNA regulatory elements in the mature transcript by inducing an alternative polyadenylation event.

To ascertain whether increases in CstF-64 protein cause a poly(A) site switch in LPS-induced genes as demonstrated for our reporter mini-gene, we found, through a PCR-based approach, that the abundance of transcripts that result from distal poly(A) site usage for Id-2 and MMP-9 are reduced under LPS-stimulating and CstF-64 over-expressing conditions (Figure 11B). Similarly, the use of the distal poly(A) site on the CTE-1 transcript is reduced under LPS-stimulating conditions (Figure 11B). Because the over-all expression of CTE-1 is greatly reduced under CstF-64 over-expressing conditions (Table 2 and Figure 8), we could not ascertain whether a shift in poly(A) site choice occurred or that the amount of transcript generated from the distal poly(A) site is below our level of detection. A more sensitive technique such as Northern Blot or RNase protection assay will need to be performed to determine the effects of CstF-64 over-expression on the CTE-1 transcript. Interestingly, the same technique used to

demonstrate a decrease in distal poly(A) site use on Id-2, CTE-1 and MMP-9 transcripts showed a definitive increase in distal poly(A) site selection on Tyki and Fabp-4 transcripts under both LPS-stimulating and CstF-64 over-expressing conditions (Figure 11C).

Showing a change in the long form of the transcript under CstF-64 over-expression while also showing a change in overall gene production (Figures 8 and 11) proves that an alternative poly(A) site choice event does occur on these five genes and likely contributes to the measured change in gene expression. Closer analysis of the nucleotide sequence that lies between these genes' poly(A) sites shows the presence of a variety of putative AREs (Figure 12). Functional AREs can have a variety of permutations, most of which contain the core element AUUUA and may have additional adenosine and/or uridine residues flanking it [Table 5 and reviewed in (150)]. Non-AUUUA containing U-rich regions have also been described [Table 5 and reviewed in (150)]. For Id-2, there are two AREs that possess the consensus core element AUUUA and a non-AUUUA containing U-rich element just 3' to the proximal poly(A) site (Figure 12). Use of any of the three distal poly(A) sites would result in the retention of these AREs to the final gene transcript. Subsequently, preferred use of the proximal poly(A) site, as observed under LPS-stimulating and CstF-64 over-expressing conditions would result in a transcript that lacks these putative regulatory elements (Figure 12). For CTE-1, there are two basal ARE core elements just 5' to the distal poly(A) site. Most interesting is the extensive U-rich element that lies just downstream of three promoter proximal poly(A) signals, any of which could associate with the indicated GU-rich region to create a functional poly(A) site (Figure 12). The U-rich element spans over 180 nucleotides and is 60% U-rich. However, unlike the U-rich mRNA instability regions defined in the 3'-UTR of c-*fos* and c-*jun* that are also rich in adenosine residues (20, 107), the remaining 40% of the U-rich region is composed primarily of cytidines (Figure 12).

Because the removal of this CU-rich region correlates with a reduction in gene expression, this element might in fact be a novel mRNA stability element that bestows a greater half-life to the longer CTE-1 transcripts. MMP-9 possesses two AUUUA-containing AREs between its two poly(A) sites. Thus, preferred use of the proximal poly(A) site observed under LPS-stimulating and CstF-64 over-expressing conditions would remove these putative regulatory elements from the final gene product (Figure 12). If these putative AREs are found to influence mRNA half-life, this would explain how simply over-expressing CstF-64 in the RAW macrophages would increase the amount of final gene product. Because the sequence requirements for an ARE to be functional are diverse, the authenticity of these putative *cis*-elements would have to be studied for each gene.

Surprisingly, we observed an increase in distal poly(A) site selection under LPS stimulation and CstF-64 over-expression for both Tyki and Fabp-4 (Figure 11C). An increased selection of the distal poly(A) site of Tyki would incorporate two core element-containing AREs into the final transcript (Figure 12). Likewise, distal poly(A) site usage in Fabp-4 would incorporate a fairly large U-rich element and a core ARE element with an associated stretch of uridines (Figure 12). Because the gene expression of both Tyki and Fabp-4 are increased under LPS-stimulating and CstF-64 over-expressing conditions, the AREs of these genes that are retained in the long form of the transcript may display mRNA stabilizing characteristics. The increase in distal poly(A)site usage observed for Tyki and Fabp-4 is the opposite of what we observed on our report mini-gene (Figure 5) and for Id-2, CTE-1 and MMP-9 gene transcripts (Figure 11B), and counterintuitive to our idea of increases in CstF-64 expression driving proximal poly(A) site use. One possibility for the increased use of the distal poly(A) site under LPS-stimulating and CstF-64 over-expressing conditions would be if one of the proteins

identified to increase in expression under these conditions (Table 2) binds to the mRNA near the proximal poly(A) site and prevents its use. One candidate for this ascribed function is Hermes, a member of the largest family of RNA-binding proteins, a factor that has demonstrated a role in *Xenopus laevis* oocyte development (155) and is highly expressed in the myocardium of the developing heart (53). Although the exact role Hermes has on RNA regulation is not known, researchers suspect that Hermes plays a role in regulating translation and/or mRNA stability in order to fine-tune the genetic program of developing tissue. Further experiments exploring the RNA-binding specificity of Hermes and its possible gene-specific interaction to Tyki and Fabp-4 transcripts would be necessary to test this hypothesis.

However, a gene need only possess a single poly(A) site for the cleavage step in pre-mRNA processing to have an impact on gene transcript production. Because of the diversity in sequence of the *cis*-elements both upstream and downstream of the cleavage site (9, 79), the cleavage step of pre-mRNA processing is sensitive to the amount of proteins that bind to these elements. That is, increases in either CPSF-160 or CstF-64, as we have shown here, can theoretically increase the formation of the cleavage/polyadenylation complex at any particular poly(A) site, especially those that deviate from the consensus sequence parameters. Furthermore, due to the growing evidence that the efficiency of 3'-end cleavage of pre-mRNA is also influenced by auxiliary upstream and downstream elements that lie near the poly(A) signal and GU-rich elements (79, 98), the amounts of *trans*-acting factors that bind to these regions may also influence the cleavage reaction. Indeed, hnRNP F and hnRNP H/H' are factors that bind to G-rich elements in RNA (19), have been found within the cleavage/polyadenylation complex (146), and whose expression levels can influence pre-mRNA processing (5, 146). Regardless of the mechanism, an increase in 3'-end processing of pre-mRNA transcripts that

contain a single poly(A) site would produce more mature mRNA to be template for protein production.

Through over-expressing CstF-64 in RAW 264.7 macrophages, we have demonstrated for the first time that altered expression levels of a single mRNA processing factor can influence the expression of multiple genes. We showed that the mechanism of mRNA 3'-end processing contributes in part to the genetic profile changes observed upon macrophage stimulation with LPS, emphasizing the impact post-transcriptional 3'-end processing has on gene expression in macrophages stimulated by gram-negative bacteria.

5. AIM 2: Differential Phosphorylation of RNAP-II Carboxy-Terminal Domain in A20 B-Lymphoma Cells and AxJ Plasma Cells

5.1. ABSTRACT

In order for plasma cells to secrete immunoglobulin (Ig), a poly(A) site switch needs to occur within the 3'-UTR of the Ig heavy chain gene that would result in the exclusion of membrane spanning exons and the production of the secretory form of Ig. This alternative polyadenylation event results from an increase in selection of a weak, promoter-proximal poly(A) site correlated with an increase in the binding activity of CstF-64, a protein involved in eukaryotic pre-mRNA 3'-end processing. With the evidence that multiple pre-mRNA processing factors are recruited to the RNAP-II through the phosphorylation of the carboxy-terminal domain (CTD), we hypothesized that there is a change in the phosphorylation state of RNAP-II CTD between A20 memory B-cells and AxJ Ig-secreting plasma cells that contributes to this poly(A) site switch. To obtain RNAP-II CTD phosphorylation data at the Ig heavy chain promoter, we sequenced the IgG2a promoter shared by both A20 and AxJ cells using a 5'-RACE PCR approach. By employing the chromatin immunoprecipitation (ChIP) technique across the IgG2a heavy chain gene, we detected an increase in both Serine-2 and Serine-5 phosphorylation of RNAP-II CTD at the promoter and variable regions of AxJ cells over A20 cells. We also detected a significant reduction in Serine-2 and Serine-5 phosphorylation at the Ig heavy chain intronic enhancer (Eμ) in both cell lines. Levels of Serine-2 and Serine-5 phosphorylation beyond the Eμ were low and unchanged through the remaining portion of the genomic locus in both cell lines. From these results, we conclude that the phosphorylation level of RNAP-II CTD that are different at the promoter and variable regions of the IgG2a heavy chain gene between A20 memory cells and AxJ plasma cells indicates either an increase in polymerase loading at the

promoter in AxJ plasma cells or a change in the phosphorylation state of RNAP-II CTD. Experiments to test these hypotheses are outlined and interpretations of the possible results are discussed.

5.2. INTRODUCTION

5.2.1. Transcription initiation

RNA polymerase type 2 (RNAP-II) is the polymerase responsible for the transcription of the vast majority of eukaryotic genes. RNAP-II is recruited to the TATA box of Pol-II promoters by the multi-subunit general transcription factor TFIID. Recruitment of TFIID to the TATA box is facilitated by one of its subunits, TATA binding protein (TBP). Sequential recruitment of the TFIIA, TFIIB, TFIIE, TFIIF, and TFIIH general TFs and RNAP-II follow thereafter forming the pre-initiation complex of transcription [reviewed in (30)]. To initiate promoter release of the polymerase and begin transcription initiation, the kinase activity of CDK7, CDK8 and CDK9, subunits of TFIIH, phosphorylates the carboxy-terminal domain (CTD) of RNAP-II (83, 114).

5.2.2. RNA Polymerase II Carboxy-Terminal Domain

The CTD of RNAP-II consists of a conserved heptad amino acid stretch repeated 26 times in budding yeast and 52 times in vertebrates and whose sequence is conserved across many eukaryotic species [(29) and reviewed in (105)]. The repeat amino acid sequence YSPTSPS possesses five residues that are phosphorylation acceptors; however Serine-2 and Serine-5 are the residues most commonly phosphorylated *in vivo* (156). Because of this, RNAP-II purified from mammalian cell extracts has been observed in several forms based on the phosphorylation status of the CTD. A hypophosphorylated form of RNAP-II CTD is associated with the pre-initiation complex and designated RNAP-IIA (128). Subsequently, the hyperphosphorylated form of RNAP-II CTD, designated RNAP-IIO, is associated with the release of the polymerase from the promoter and the beginning of transcription (128). Notably, other intermediately phosphorylated forms of RNAP-II have been observed and correspond to an intermediate level

of phosphorylation. Indeed, the precise pattern of phosphorylation of the CTD of RNAP-II plays a significant role in the initiation of transcription and co-transcriptional mRNA processing (see below).

5.2.3. Kinases that phosphorylate the CTD of RNAP-II

CTD phosphorylation has been shown to be required for promoter clearance and transcription initiation and elongation (85, 154). A large variety of kinases have been reported that are responsible for the post-translational modifications of the CTD of RNAP-II and include cell-cycle dependent kinases (CDK) and map kinases (MAPK) (83, 103). The activity and specificity of many of these kinases are highly regulated and are often coordinated with the cell cycle (83). Furthermore, the specificity of these kinases is precise in that the particular residues within the CTD of RNAP II that are phosphorylated are kinase dependent and are affected by the phosphorylation state of neighboring residues (118, 119, 141). As an added level of specificity control, the association of transcription factors can further dictate the precise CTD residues that are phosphorylated. For example, CDK7 has been shown to phosphorylate both serine-2 and serine-5 residues. However, the combinatorial effects of CDK7 and TFIIE enforce the preference of serine-5 phosphorylation (122). As CTD phosphorylation is a highly regulated event, the removal of phosphates from the CTD by phosphatases is also regulated, however much less is known about these events [reviewed in (105)].

5.2.4. Early links between mRNA processing and transcription

One of the first pieces of evidence that transcription and mRNA processing are linked events was the detection of CPSF subunits during the purification of TFIID-associated proteins (32). Subsequent reconstitution of an inducible, *in vitro* transcription system demonstrated that CPSF and TFIID are co-localized to the promoter. However, upon chemical induction of

transcription, TFIID remained at the promoter while CPSF migrated with the polymerase (32). Another early study had demonstrated that the 3'-end processing reaction is inhibited when the CTD of RNAP II is truncated (89). In fact, reconstitution experiments demonstrated that RNAP II CTD is a required element for *in vitro* pre-mRNA cleavage reactions and that the addition of the hyperphosphorylated form of CTD (RNAP-IIO) provided more cleavage activity to the reaction than the hypophosphorylated form (RNAP-IIA) (65). Furthermore, both CPSF and CstF proteins were shown to bind CTD affinity columns and were co-purified in RNAP II-containing high-molecular-mass complexes (89). Lastly, immunohistochemistry has co-localized subunits from both CPSF and CstF to nuclear domains closely associated with actively transcribed genes and newly synthesized RNA (126); this association occurs during cell cycle progression (127).

5.2.4.1. The Connection between RNAP-II CTD phosphorylation and mRNA processing

The CTD of RNAP-II has long been thought to behave as a landing platform for a variety of mRNA processing factors involved in the capping, splicing and 3'-end cleavage of nascent pre-mRNA. Because the processing of mRNA is reduced concordantly with the prevention of CTD phosphorylation, the recruitment of these mRNA processing factors to the CTD has been shown to be phosphorylation dependent (12). Crystallographic studies have shown that the phosphorylation of serine-2 within the heptad repeat stabilizes a β-turn that interacts with the CTD-interacting domains of 3'-end processing factors (91). Biochemical studies have elucidated specific regions of the RNAP II CTD that are charged with different mRNA processing mechanisms. By engineering truncated forms of the RNAP II CTD, it was discovered that the carboxy terminus of the CTD supports all three major pre-mRNA processing events, while the

amino terminus of the CTD only supports capping (47). A closer inspection of the ten unique amino acid residues that are located 3' to the terminal heptad repeat show that this sequence also significantly supports splicing and 3'-end processing (48). Interestingly, the phosphorylation state of this 10-residue sequence does not appear to be necessary for its role in pre-mRNA processing (48).

5.2.4.2. Differential phosphorylation of RNAP-II CTD and the recruitment of mRNA processing factors

Although the conserved heptad repeat of the carboxy-terminus of RNAP-II possesses five phosphorylation accepting amino acid residues, the main residues targeted by CTD kinases are Serine-2 and Serine-5 (156). Because of the necessity of CTD phosphorylation for promoter clearance and transcription elongation, the fact that pre-mRNA processing is dependent on CTD phosphorylation, and that pre-mRNA processing factors have been shown to bind to phosphorylated CTD, it was questioned whether pre-mRNA processing factors were associated to the transcriptional machinery throughout transcription or recruited in a phosphorylation-specific and location-specific manner across the gene. Employing the chromatin immunoprecipitation (ChIP) technique, researchers have been able to detect the phosphorylation state of the RNAP-II CTD on Serine-2 and Serine-5 across particular yeast (2, 73, 75, 125) and mammalian (23) genes *in vivo*. Furthermore, this technique has been expanded to identify pre-mRNA processing factors that are directly associated with the polymerase provided an appropriate immunoprecipitation-competent antibody for the protein of interest is available.

Using ChIP, the association of yeast pre-mRNA capping enzymes was shown to occur at the 5' of RNAP-II transcribed genes and was concomitant with Serine-5 phosphorylation of the RNAP-II CTD (75, 125). Furthermore, both the association of the capping enzymes and the

Serine-5 phosphorylation state of RNAP-II CTD was reduced as the transcription machinery migrated toward the 3'-end of the gene (125). Interestingly, as Serine-5 phosphorylation was reduced in the coding regions of yeast genes, it was shown that the Serine-2 phosphorylation state of RNAP-II CTD increased as the transcriptional machinery migrated toward the 3'-end of the gene (75). Coupled to the increase in Serine-2 phosphorylation along the coding region of yeast genes was the association of RNAP-II elongation factors (113) and pre-mRNA 3'-end processing factors (2, 77, 81).

Consistent with what was observed in yeast, Serine-5 phosphorylation was strong at the promoter of the mammalian DHFR and γ-actin genes and was subsequently reduced as the transcriptional machinery migrated across the gene (23). However, in contrast to what is typically observed on yeast genes, the Serine-2 phosphorylation of RNAP-II CTD was distributed more evenly across the genes (23). Furthermore, Cheng *et al*. demonstrated that the high level of RNAP-II measured at the promoter of these genes, which was reduced within the open reading frame, was due to the RNAP-II falling off the gene and incomplete transcription (23). The association of mammalian pre-mRNA processing factors to the transcriptional machinery and the correlation to the phosphorylation state of the RNAP-II CTD is currently unknown.

5.2.5. Maturation of B-lymphocytes and Ig Secretion

5.2.5.1. B-lymphocyte Biology

B lymphocytes are a part of the humoral immune system and are generated in the bone marrow. Each B-cell expresses a B-cell receptor (BCR) in the form of membrane-bound immunoglobulin unique to a particular antigen. The maturation of naïve B-lymphocytes into

immunoglobulin (Ig) secreting plasma cells (PCs) occurs as a result of B-cell recruitment to germinal centers in the spleen and lymph nodes and subsequent activation following exposure to soluble antigen or antigen associated with follicular dendritic cells. During this phenotypic change, the immunoglobulin gene undergoes several rounds of somatic hypermutation which result in slight changes in antigen specificity. If the change in specificity to its specific antigen is lessened, this will result in B-cell anergy. On the contrary, an increase in affinity for the antigen results in further maturation and differentiation. Naïve B-cells begin by expressing surface bound IgM. Along with B-cell differentiation, the Ig gene undergoes isotype switching with the aid of T-helper cells to IgG, IgA or IgE expressing B-cells.

5.2.5.2. Immunoglobulin Germline Configuration

The murine germline configuration of the immunoglobulin heavy chain gene includes exons that code for hundreds of variable regions, twenty-three diversity regions, four joint regions and clusters of constant region exons specific for a particular isotype (Figure 14). Through the B-cell developmental process, the heavy chain Ig gene undergoes sequential germline recombination first between the diversity (D) and joint (J) regions creating a DJ gene segment. Subsequently, recombination then occurs between the variable (V) region and the DJ gene segment creating a final VDJ Ig gene segment. Each variable region is associated with its own leader sequence (L) and promoter (Figure 14).

Located downstream of the VDJ gene segment is the intronic Ig enhancer, Eμ (Figure 14), that possesses several conserved *cis*-elements called E-boxes that bind to E-box binding proteins. Activation of the intronic enhancer by these E-box binding proteins is necessary for proper B-cell development (55) and contributes to Ig gene transcription (149). Further downstream of the intronic enhancer are the constant region exons (Figure 14). In the naïve B-

cell germline configuration, the mu constant region exon cluster (Cμ) lies immediately downstream from the VDJ gene segment, and thus results in IgM production in naïve B-cells. Upon isotype switching, another genetic recombination even occurs between switch signals that precede each constant region exon cluster, thus relocating constant region exons from other isotypes immediately downstream of the VDJ gene segment.

5.2.5.3. Alternative Ig mRNA processing in the plasma cell

A hallmark of B-cell differentiation into a PC is the ability of the PC to secrete copious amounts of antigen-specific antibody. In order for Ig to be secreted, an alternative post-transcriptional processing event must occur that results in the selective use of the upstream secretory poly(A) site (pAs) in the Ig gene (Figure 15). This selection in upstream poly(A) site use would preclude the incorporation of membrane spanning exons to the final Ig transcript. Several hypotheses have emerged and been tested to explain the shift in poly(A) site usage on the Ig gene in plasma cells.

Because the selection of the upstream pAs removes the possibility of incorporating the membrane spanning exons into the mature Ig transcript, which necessitates the selection of a 5'-splice site that is embedded in the exon immediately upstream of pAs, it is possible that changes in splicing and/or polyadenylation activity in the cellular environment can tip the balances creating the secretory form over membrane form of Ig. Previous studies have demonstrated an increase in polyadenylation efficiency in PCs compared to B-cells without a change in splicing efficiency on the IgM gene (110). However, further studies have shown that the particular cellular environment can determine whether splicing or cleavage/polyadenylation is the dominant processing event (15). Work from our lab has shown that regardless of sequence

composition, an increased use of a weak upstream poly(A) site over a stronger distal poly(A) site is observed in PCs compared to naïve or memory B lymphocytes (87).

Alternatively, there could be an increase in the early termination of Ig transcription facilitated by an increase in termination factors or a decrease in transcription elongation factors in PCs. A reduction in the full extension of the Ig transcript past the membrane poly(A) site (pAm) would result in an increase in pAs usage This had been shown to occur on the IgM gene in B-cells exhibiting a PC phenotype (148). However, studies from our lab have shown that transcription continues far past the pAm on the IgG2a and IgG2b heavy chain genes in memory B-cells (92). Aside from transcription termination, it has been suggested that transcriptional pausing downstream of the pAs in plasma cells may be a contributing mechanism to increased pAs selection. Indeed the removal of nucleotide sequences immediately downstream of the IgM and IgA pAs decreased the production of secretory IgM in PCs (109).

5.2.5.4. Secretory Ig Production and CstF-64

The selection of the upstream weak poly(A) site in the Ig gene is essential for the production of the secretory form of Ig. LPS stimulation of primary splenic B-cells had demonstrated an accumulation of CstF-64 protein that correlated with an increase in secretory Ig mRNA production (139). Furthermore, over-expression of CstF-64 in chicken DT-40 B-lymphoma cells caused an increase in secretory Ig mRNA production (139). However, studies from our lab have shown an increase in the binding efficiency of CstF-64 in plasma cells in the absence of increases in protein production (44). Furthermore, the increased ability of CstF-64 to bind to RNA substrates in the absence of increases in protein levels have been demonstrated in the context of viral infections (84, 90). One explanation for the increase in CstF-64 observed in the LPS-stimulated splenic B-cells is that LPS induces rapid proliferation of B-cells (117) and

our lab has shown that CstF-64 protein levels increase when cells enter the Go-S portion of the cell cycle (86). While there is no denying that increased levels of CstF-64 protein cause the increase in selection of weaker poly(A) sites, particularly when CstF-64 is the limiting constituent to CstF trimer formation as demonstrated through immunoprecipitation studies (86), we believe that increases in weak poly(A) site selection can occur on the Ig gene in the absence of CstF-64 protein increase; perhaps by an increase in the local concentration of the steady-state level of CstF at the site of pre-mRNA cleavage or increased recruitment of CstF to the migrating polymerase on the Ig heavy chain gene.

5.2.6. The Strength of our Experimental System Through the Use of the AxJ Hybridoma Cell Line

When using ChIP to illustrate the dynamics of the phosphorylation state of RNAP-II CTD and how that correlates to transcription initiation, elongation and the association of pre-mRNA processing factors to the CTD of RNAP-II across a gene's genomic sequence, it is important to obtain a complete picture that includes data from one end of the gene to the other. Therefore it is necessary to have the promoter sequence of the gene of interest for primer design, particularly since RNAP-II is recruited to the promoter and some pre-mRNA processing factors have demonstrated preferential association to the polymerase while at the promoter. Thus, we have employed a 5'-RACE PCR approach to obtain the promoter region of the A20 active IgG2a locus. Because our studies are focusing not only on the dynamics of RNAP-II CTD phosphorylation across the Ig genomic locus in A20 memory B-cells, but also on the difference in the phosphorylation pattern and association of pre-mRNA processing factors between memory B-cells and Ig-secreting plasma cells, we need to be able to perform ChIP on the identical Ig promoter and V-region from both membrane Ig generating B-cells and secretory Ig generating

plasma cells. However, we have been unable to induce the A20 memory B-cell line to secrete IgG2a through LPS or cytokine stimulation (data not shown). Furthermore, we could not simply choose a murine plasma cell line that expresses the same Ig isotype as the A20 memory cells (IgG2a), for that would only supply a cell line expressing the same constant regions with a different promoter and uniquely recombined VDJ region. Therefore, our lab had generated a hybridoma cell line through the fusion of A20 memory B-cells and a subclone of the J558 plasma cell line that lacks heavy chain (J558L) (93). This AxJ hybridoma expresses the identical heavy chain gene (including promoter and variable regions) as the A20 memory cell line, demonstrates preferential use of the secretory poly(A) site on the Ig gene (93), secretes Ig (Figure 16) and displays a genetic profile similar to primary plasma cells (data not shown). With this hybridoma unique to our laboratory, we have the ability to perform ChIP on the identical Ig heavy chain genomic locus between a memory B-cell line that generates primarily membrane-bound Ig and a plasma cell line that generates primarily secretory Ig.

5.3. MATERIALS AND METHODS

5.3.1. A20 IgG2a promoter sequencing by 5'-RACE PCR

The Universal GenomeWalker™ Kit from BD Biosciences (K1807-1) was used to obtain the promoter region for A20 IgG2a. Beginning with the published sequence for the variable region of A20 B-lymphoma cell line (39), two gene-specific anti-sense primers (GSPs) were created that targeted this known sequence (Figure 17 and Table 6). In brief, 2.5 μg of A20 genomic DNA, purified using the DNeasy kit (Qiagen), was digested with one of four blunt-end DNA restriction enzymes (DraI, EcoRV, PvuII and StuI) provided with the Universal GenomeWalker™ Kit creating four genomic DNA libraries. After observing a streak of DNA on an ethidium bromide-stained agarose gel from a fraction of the digestion reaction confirming successful DNA digestion, each digestion was purified and ligated to the GenomeWalker™ adaptors according to manufacturer's instructions. Two sequential PCR reactions were subsequently performed using Adaptor Primer 1 (AP1) and GSP1 followed by Adaptor Primer 2 (AP2) and GSP2 (Figure 17A**)**. Specifically, 1 μl of each purified genomic DNA library was combined with 10 pmol each of forward AP and reverse GSP, Mg^{2+}, dNTP mix, 10X reaction buffer and Advantage Genomic Polymerase Mix according to manufacturer's instructions. A modified touch-down PCR program suggested by the protocol was employed to reduce non-specific amplification. A second round of touch-down PCR was performed using 5 μl of first-round PCR product stemming from each genomic DNA library, and 10 pmol each of forward AP2 and reverse GSP2. A significant PCR product was observed from the two sequential rounds of PCR stemming from the PvuII-created genomic DNA library. This PCR product was gel purified using the QIAEX II gel extraction kit (Qiagen, 20021) and ligated to the pCR2.1 plasmid as part of the TA Cloning® Kit (Invitrogen, 25-0024) according to manufacturer's

instructions. The ligated plasmid was amplified in One Shot® DH5α-T1 competent cells (Invitrogen, 12297-016), purified over a Maxi-prep column (Qiagen) and sequenced (University of Pittsburgh, School of Med., Biotechnology Center) using M13 sense and antisense primers that bind to pCR2.1 regions that flank the inserted PCR amplicon (Table 6).

5.3.2. IgG2a ELISA

The A20 memory B-cells (ATCC) and AxJ plasma cell hybridomas (93) were maintained in IMDM with 5% horse serum, 10 mM HEPES, and 0.1% penicillin-streptomycin. Two million A20 and AxJ cells were seeded in 4 ml media for each time point in 6-well dishes. Aliquots of media were retained at 2, 4, 6, 8, 10 and 12 hours. Goat anti-mouse IgG antibody (Sigma M8642) was diluted in PBS and used to coat a 96-well ELISA plate (DYNEX, 1010) with 200 ng capture antibody per well and incubated at 4°C overnight. The wells were subsequently washed 3 times with PBS/T, blocked with 10% BSA/PBS/T for 2 hours at room temperature, and then washed once with PBS/T. One hundred microliters of sample was applied in triplicate to each well and incubated at 37°C for one hour. The wells were then washed five times with PBS/T followed by applying AP-conjugated goat anti-mouse antibody (Zymed, 62-6622) diluted 1:1000 in 1%BSA/PBS/T to each well and incubated for two hours at room temperature. The wells were then washed five times with PBS/T and 100 μl of room temperature p-NNP (Chemicon, ES009) was applied to each well for 15 minutes room temperature with gentle rotation followed by colorimetric measurement at 405 nm.

5.3.3. Chromatin Immunoprecipitation

For crosslinking, $5x10^7$ A20 cells or AxJ cells were chemically fixed with the addition of formaldehyde (1% final concentration) directly to the cell culture suspension for ten minutes at room temperature while rocking. Fixation was stopped by the addition of glycine at a final

concentration of 125 mM for ten minutes at room temperature while rocking. Cells were then pelleted, washed with PBS, pelleted again and stored at -80°C until needed. The ChIP procedure was performed using the ChIP-IT™ kit (Active Motif) as per manufacturer's instructions with minor exceptions noted below. In brief, the nuclei from $5x10^7$ cells were harvested using cold ChIP-IT™ Lysis Buffer and protease inhibitors (PI) on ice for 30 minutes followed by dounce homogenization and subsequent centrifugation. Nuclei were resuspended in ChIP-IT™ Shearing Buffer and split into three equal volumes in eppendorf tubes. Chromatin was fragmented to 1000 base pairs or less using a Microson XL 2000 (Misonix) for 20 15-second pulses at level 7 with the tubes sitting in an ice water/salt slurry, allowing at least 30 seconds to lapse between pulses. Each new chromatin preparation was tested for shearing efficiency and DNA concentration as per ChIP-IT™ protocol.

Ten micrograms of fragmented chromatin DNA were used for each chromatin immunoprecipitation. Chromatin was pre-cleared using 100 µl supplied protein G beads, ChIP IP Buffer and PI to a total volume of 210 µl per 10 µg chromatin DNA. Pre-clearing was performed for 1 hour at 4°C with gentle rocking. Ten µl of pre-cleared chromatin was combined with 90 µl water for the Total Input sample and stored at -20°C until needed. The remaining volume of pre-cleared chromatin was evenly distributed into as many tubes as the number of planned immunoprecipitations (10 µg per IP). Antibodies were added to the pre-cleared chromatin and the ChIP reaction was incubated overnight at 4°C with gentle rocking. One hundred µl of Protein G resin (ChIP IT™ kit) or anti-IgM-specific agarose (Sigma, A4540) was added to each ChIP reaction and rotated at 4°C for 90 minutes followed by collection by centrifugation. The antibody/chromatin-resin complex was subject to one 400 µl wash using ChIP IT™ IP Buffer with PI, four 400 µl washes using ChIP IT™ Wash Buffer 1 with PI, one

400 µl wash using ChIP IT™ Wash Buffer 2 with PI and two 400 µl washes using ChIP IT™ Wash Buffer 3 (no PI). All washes were carried out with cold wash buffers at room temperature. Elution of antibody/chromatin complexes from the resin was performed using 50 µl of ChIP IT™ Elution Buffer incubated at room temperature for 15 minutes, flicking tubes every 3 minutes. The resin was collected by centrifugation, the eluate was retained in a new tube and the elution process was repeated once. The eluates were combined and reverse crosslinking of the protein from the DNA was performed using 6.5 µl of 5M NaCl (325 mM) and 1 µl RNase A at 67°C overnight. Proteins were subsequently digested using 2 µl 0.5 M EDTA, 2 µl 1 M Tris-Cl pH 6.5 and 2 µl Proteinase K solution (ChIP IT™ kit) and incubated at 42°C for 2 hours. Purification of the immunoprecipitated DNA was performed over the DNA purification columns contained within the ChIP IT™ kit as per manufacturer's instructions. Typically, two to three volumes of nuclease-free water was added to the eluted DNA of all samples in order to expand the sample volume for subsequent QPCR analysis. Antibodies were obtained from the following sources and the amounts used per ChIP are indicated: anti-TFIIB positive control (ChIP IT™, 10 µl or 4 µg), purified mouse IgG negative control (ChIP IT™, 10 µl or 4 µg), anti-Serine-2 phosphorylated RNAP-II CTD (H5, Covance; 10 µl or ~25 µg), anti-Serine-5 phosphorylated RNAP-II CTD (H14, Covance; 15 µl or ~60 µg), purified mouse IgM negative control (ICM Pharmaceuticals, 20 µl or 25 µg).

5.3.4. Real-Time PCR and Data Analysis

Quantities of immunoprecipitated DNA were measured using the ABI Prism 7900HT and Sequence Detection Software v2.1 (SDS 2.1) provided by The University of Pittsburgh Genomics and Proteomics Core Laboratories (GPCL) (http://www.genetics.pitt.edu/). In brief, 5 µl of immunoprecipitated DNA were combined with 2X SYBR Green PCR Master Mix (Applied

Biosystems) and 6.25 nmoles each forward and reverse primers to a final reaction volume of 25 µl. The following PCR program was used for all ChIP DNA samples: Taq activation for 12 minutes at 95°C, followed by 40 cycles of a 15 second denaturing step at 95°C and a 1 minute annealing step at 60°C. Quantitation of immunoprecipitated DNA was normalized at every genomic locus by comparing its signal to that obtained from a 1:10 dilution of the total input DNA sample using the same primer pair and calculating the "% input" using the following formula: % input = $2^{\Delta CT}$x10 where ΔCT is the difference in CT values between the diluted input DNA and the IP'd DNA at a chosen threshold value. ChIPs were performed between two and four times per locus from at least two different sources of chromatin. Averages and SEM were calculated using Excel from Microsoft Office 2000.

5.4. RESULTS

5.4.1. Sequencing the A20 IgG2a promoter

By consulting the published V-region sequence of the A20 B-cell lymphoma line (39) and the genomic sequence for the murine IgG2a constant regions (Accession J100470) and membrane spanning exons and poly(A) sites (Accession J100471), we designed primers that spanned the approximately 12 kb of genomic DNA (Figure 18). However, a great deal of information can be obtained by performing ChIP experiments on the promoter region of IgG2a, and this can only be done by determining the sequence of the promoter. Because the sequence information obtained for the A20 Ig variable region was generated from cDNA, the promoter region contained within the 5-UTR was not available for sequencing. The A20 B-cell line is a clonal population of memory B-cells where the identical VDJ recombined IgG2a gene and promoter are active in every cell. Therefore, through a 5'-RACE PCR approach upstream from the known variable region of A20 IgG2a, we set out to obtain the promoter sequence. The VDJ recombination event involves the J2 segment of the J-locus (39), so we designed the first gene-specific primer (GSP1) to target the genomic region between J2 and J3 (Figure 17A). The target sequence for the second GSP (GSP2) spanned the region where the V, D and J-regions have uniquely joined in A20 B-cells through recombination (Figure 17A).

Shown in Figure 17 (Figure 17B and 17C**)** is the promoter sequence obtained from the 5'-RACE PCR protocol. The Ig heavy chain gene promoter contains multiple conserved sequence motifs that bind B-cell specific transcription factors and are required for Ig production. Octamer binding proteins bind the Octamer element (ATGCAAAT) that lies approximately fifty nucleotides upstream from the transcription start site of Ig variable regions and is a strong contributor to Ig gene expression (8, 58). A heptamer motif is often associated with the Octamer

element in Ig promoters and has been shown to both bind Octamer binding proteins and contribute to Ig transcription (112). Both of these elements appear in the A20 V-region promoter (Figure 17C). Also located in the A20 V-region promoter are a C/EBP transcription factor binding site, a pyrimidine stretch and a TATA box (Figure 17C), all of which have been shown to be functionally important to Ig transcription (28, 42).

5.4.2. Comparing the phosphorylation state of RNAP-II CTD across the IgG2a genomic locus between A20 memory B–cells and AxJ plasma cells

Much work has been published describing the association of pre-mRNA processing factors to the CTD of RNAP-II (32, 47, 80, 89, 91, 113) and how these interactions are CTD phosphorylation dependent (2, 73, 75, 81, 125). Because there is conflicting evidence on whether the alternative poly(A) site selection that occurs on the Ig gene in plasma cells is dependent on CstF-64 protein increases (86, 139), we hypothesize that the association of CstF-64 to the transcriptional machinery is increased in plasma cells resulting in an elevated local concentration of CstF-64 to the Ig transcript that may or may not be observed by Western Blots on whole cell lysates or nuclear extracts. To test this, we first set out to determine if there are phosphorylation changes on the RNAP-II CTD across the Ig genomic locus and if the pattern changes between A20 memory cells and AxJ Ig-secreting plasma cells.

5.4.2.1. Serine-2 phosphorylation of RNAP-II CTD in A20 memory B-cells and AxJ plasma cells

By employing chromatin immunoprecipitation (ChIP), we were able to assess the phosphorylation state of RNAP-II CTD across the Ig genomic locus. To do this, we designed several primer pairs that span the Ig genomic locus from the promoter region to downstream of the secretory poly(A) site (Figure 18). Because our shearing conditions aimed at achieving

genomic DNA fragments less than 1000 base pairs, our primer pairs were designed with at least 1000 base pairs of genomic sequence between them to assure that the signals we achieved were authentic to the region of the Ig targeted (Figure 18). Using an antibody specific to the Serine-2 phosphorylated form of RNAP-II CTD, we measured a decrease in Serine-2 CTD phosphorylation from the A20 Ig heavy chain promoter and variable regions to the intronic enhancer (Eμ) (Figure 19A). Loci 3' of the Eμ showed phosphorylation levels lower than that which was measured at the A20 Ig heavy chain promoter and variable regions. The level of Serine-2 CTD phosphorylation on the AxJ Ig promoter was approximately 6.5-fold more than that observed in A20 memory B-cells. Furthermore, Serine-2 CTD phosphorylation at the AxJ variable region was approximately 5-fold more than what was measured at the A20 Ig promoter (Figure 19A). Similar to the dynamics of Serine-2 CTD phosphorylation observed in A20 cells, the phosphorylation state of RNAP-II CTD at the Eμ was reduced to very low levels. Interestingly, the Serine-2 CTD phosphorylation increased slightly 3' of the Eμ, but did not rebound to the level observed at the AxJ promoter and variable regions and appeared similar to that which was measured for A20 Ig gene in profile (Figure 19A**)**.

5.4.2.2. Serine-5 phosphorylation of RNAP-II CTD in A20 memory B-cells and AxJ plasma cells

Through the use of an antibody specific to the Serine-5 phosphorylated form of RNAP-II CTD, we observed that the overall profile of Serine-5 phosphorylation in A20 memory B-cells resembles that observed for Serine-2 phosphorylation in the same cell type. Furthermore, the profile of Serine-5 phosphorylation in AxJ plasma cells mirrored the Serine-2 phosphorylation profile observed in the same cell type. Specifically, Serine-5 phosphorylation in the A20 cells

was moderate at the promoter and variable regions, and then dropped to very low levels at the Eμ (Figure 19B). Downstream of the Eμ, Serine-5 phosphorylation in A20 cells rebounded to levels slightly below what was measured at the regions upstream of the Eμ at most loci (Figure 19B). It appears that Serine-5 phosphorylation spikes at the constant region in A20 cells, however this spike is not statistically higher than the levels observed at the promoter and variable regions. Similar to the observations for Serine-2 phosphorylation of RNAP-II CTD, Serine-5 phosphorylation is approximately 6-fold higher at both the promoter and variable regions of the Ig in AxJ plasma cells compared to A20 memory B-cells (Figure 19B). Likewise, the phosphorylation level of Serine-5 dropped to very low levels at the Eμ in AxJ cells and then rebounded slightly achieving levels of Serine-5 phosphorylation similar to that measured in A20 cells.

5.4.2.3. Summary of Serine-2 and Serine-5 Phosphorylation of RNAP-II CTD

From this, we conclude that the Serine-2 and Serine-5 phosphorylation of RNAP-II CTD in A20 memory B-cells is generally low across the entire Ig genomic locus, that there is a significant reduction in phosphorylation at the Eμ, and that the measured phosphorylation is higher after the Eμ to levels observed at the Ig heavy chain promoter and variable regions. We further conclude that the Serine-2 and Serine-5 phosphorylation of RNAP-II CTD at the Ig heavy chain promoter and variable regions in AxJ plasma cells is between 5- and 6-fold higher than what was measured in A20 memory B-cells, that there is a similar drop in phosphorylation at the Eμ, and that the phosphorylation levels increase downstream of the Eμ but only to that which was observed in A20 B-cells.

5.5. Conclusions and Future Directions

Our ChIP results demonstrate that the phosphorylation levels of RNAP-II CTD are dynamic across the Ig genomic locus. Specifically, the measured level of Serine-2 and Serine-5 phosphorylation of RNAP-II CTD across the Ig heavy chain genomic locus is relatively low in A20 memory B-cells. In contrast, the Serine-2 and Serine-5 phosphorylation levels of RNAP-II CTD are significantly higher at the promoter and variable regions in AxJ plasma cells, then displays similarly low levels of phosphorylation as observed in A20 memory B-cells across the remainder of the IgG2a heavy chain gene.

The results we have reported here are unlike the phosphorylation dynamics observed in actively transcribed yeast genes that exhibit high Serine-5 phosphorylation at the promoter that decrease with transcription and low Serine-2 phosphorylation at the promoter that increase with transcription (75). Our phosphorylation dynamics appear to be cell-type dependent in that both Serine-2 and Serine-5 phosphorylation share similar profiles in both A20 memory B-cells and AxJ plasma cells, yet the profiles themselves are different between the two cell lines. There was more measured phosphorylation at the promoter and variable regions of Ig heavy chain gene in AxJ plasma cells than in A20 memory B-cells. Because the phosphorylation of RNAP-II CTD is required for transcription initiation and elongation (85, 154), and because the transcription of the Ig gene in Ig-secreting plasma cells is significantly higher than naïve or memory B-cells (22), we have formulated two alternative hypotheses to explain our observations: (1) the phosphorylation level of RNAP-II CTD is higher at the promoter and variable regions of the Ig heavy chain gene in plasma cells (i.e. there are more phosphorylated residues per RNAP-II CTD molecule), or (2) the phosphorylation state of the RNAP-II CTD is unchanged, yet there is a build-up of RNAP-II

molecules at the promoter and variable regions in plasma cells associated with an increase in Ig heavy chain transcription.

By performing the ChIP technique on the Ig gene in A20 and AxJ cells using an antibody that recognizes RNAP-II regardless of the phosphorylation status of the CTD, we could answer several questions: The first, is the phosphorylation level of RNAP-II CTD measured across the Ig heavy chain genomic locus in A20 memory B-cells and AxJ plasma cells representative of the phosphorylation state of the RNAP-II CTD? If the abundance of RNAP-II across the Ig heavy chain genomic region in A20 memory B-cells and AxJ plasma cells is uniform, then the dynamics of CTD Serine-2 and Serine-5 phosphorylation observed are true in that there is little change in the phosphorylation state of RNAP-II CTD in A20 cells across the Ig heavy chain genomic locus. This would also lead us to conclude that there is a high phosphorylation state of RNAP-II CTD at the promoter and variable regions in AxJ cells that is greatly reduced at and beyond the Eμ. Furthermore, to explain the increase in RNAP-II CTD phosphorylation at the Ig heavy chain promoter in plasma cells, similar levels of RNAP-II would need to be associated to the promoter between A20 and AxJ cells.

The second question we could answer is, by measuring RNAP-II holoenzyme across the Ig heavy chain genomic region, is the increase in RNAP-II CTD phosphorylation at the Ig promoter in plasma cells related to an increase in polymerase loading? If the RNAP-II profile reveals higher levels of RNAP-II association at the Ig promoter in AxJ cells compared to A20 cells, this result would show a correlation between the increase in Ig transcription and an increase in RNAP-II promoter loading. By calculating the fold-increase in RNAP-II association in plasma cells over memory B-cells, we could estimate how much of the 6-fold increase in

phosphorylation is due to the abundance of the polymerase at the promoter and how much is due to an increase in the phosphorylation state of the polymerase.

The final question we could answer is, if there is an increase in polymerase loading at the Ig heavy chain promoter in AxJ plasma cells, what causes the decrease in detection of RNAP-II CTD phosphorylation at the Eμ? If similar levels of RNAP-II are found to associate to both the Ig heavy chain promoter and the Eμ, this would demonstrate a decrease in the phosphorylation state of RNAP-II CTD as it passes through the Eμ. If RNAP-II association drops at the Eμ as does its phosphorylation, this could indicate that either the polymerase is dissociating from the Ig heavy chain gene at this locus or that the polymerase is migrating through the Eμ at a faster rate that disallows an equal amount of RNAP-II capture by the ChIP technique. Previous studies have shown that the high level of RNAP-II association at mammalian promoters is reduced after transcription begins, due to polymerase disassociation and incomplete transcription (23). However, work from our lab has shown that there is no difference in transcription read-through of the Ig heavy chain gene between memory B-cells and plasma cells (92). Thus, if we measure a decrease in RNAP-II CTD association at the Eμ, we hypothesize that it may be due to stacking up of RNAP-II at the promoter followed by an elevated rate of transcription through the Eμ. With the transcriptional machinery traveling through the Eμ at higher speeds than at the upstream promoter and V-region, there would be less RNAP-II at any particular time at the Eμ to be captured by ChIP. Research from our lab showing an increase in the PC4 and ELL2 elongation factors in plasma cells over memory B-cells (data not shown) supports this hypothesis. However, the critical experiment using an RNAP-II specific antibody across the Ig genomic locus will need to be performed to answer the questions above. Preliminary attempts using the 8WG16 antibody that is specific for the C-terminal heptapeptide repeat present on the

largest subunit of RNAP-II and which has been previously used is eukaryotic ChIP experiments (23, 75) were unsuccessful. Troubleshooting the amount of antibody and obtaining other ChIP proven antibodies for RNAP-II are currently under way.

The main goal of this project was to identify the *in vivo* association of pre-mRNA processing factors to the transcriptional machinery across the mammalian Ig heavy chain gene. We also aimed to determine if there is a difference in association of these factors to the polymerase between memory B-cells and Ig secreting plasma cells. We hypothesize that the amount of and/or location where these factors associate with the transcriptional machinery will be different between memory B-cells and plasma cells and that these differences can be correlated to the phosphorylation pattern RNAP-II CTD exhibits across the Ig heavy chain genomic locus between these two cell lines. Unlike the majority of the yeast studies that identify the association of pre-mRNA processing factors by ChIP through the stable over-expression of a tagged version of the protein which is captured using an antibody specific to that tag, our work in the mammalian system requires us to obtain antibodies that are able to immunoprecipitate the native protein to a degree that will supply a significant signal while minimizing spurious results from non-specific binding. Our lab has developed or acquired a multitude of antibodies to mammalian pre-mRNA processing factors, some of which have been successfully used for IP studies. Through the exhaustive screening of our antibody collection, we hope to discover some antibodies to pre-mRNA processing factors, primarily ones specific for CPSF and CstF, that are compatible with our ChIP system in order to test the aforementioned hypothesis.

6. SUMMARY AND CONCLUSIONS

Post-transcriptional control of gene expression has been identified in a number of genes involved in the acute inflammatory response of the immune system. This control in gene expression is often due to changes in mRNA half-life and/or the translatability of the gene under conditions that promote this acute response and are controlled by AU-rich *cis*-elements (AREs) that lie within the 3'-UTR of the gene [reviewed in (150)]. Like other functional *cis*-elements that control transcription and 3'-end processing, AREs come in a variety of permutations that stem from a core consensus AUUUA element. Other non-AUUUA containing U-rich stretches have also been shown to regulate gene expression post-transcriptionally. There are many proteins that have been identified to bind to these regulatory *cis*-elements whose expression is also regulated under a variety of cellular conditions. Besides regulating the abundance of these *trans*-acting factors to bestow their influences through the AREs, another way to control the influence of the AREs is to dictate their presence or absence in the final gene transcript. This is done through alternative polyadenylation of the gene transcript where selection of one poly(A) site over another determines the presence or absence of the regulatory AREs in the final gene transcript.

Our work has shown that LPS-stimulation of the RAW 264.7 murine macrophage cell line causes alternative polyadenylation on a number of genes. In correlation with this measured poly(A) site switch is an increase in the level of CstF-64, a protein involved in pre-mRNA cleavage and polyadenylation. Other studies have shown that increases in CstF-64 protein induce poly(A) site switches on Ig heavy chain genes in B-cells (139), NF-ATc in effector T-cells (25) and Testis Brain RNA-Binding Protein (TB-RBP) in male germ cells (24). Despite our observation that CstF-64 protein increases in LPS-stimulated BMDMs (Figure 4), we were

unable to detect a poly(A) site switch on our reporter mini-gene in these cells due to limitations in transfecting primary cells with our reporter plasmid. However, because of our success in detecting alternative polyadenylation on the real genes Id-2, CTE-1, MMP-9 Tyki and Fabp-4 through our PCR approach (MRP-PCR) using cDNA generated from RAW macrophages, identification of these post-transcriptional poly(A) site switching events on cDNA generated from LPS-stimulated BMDMs could be easily performed and provide evidence of the post-transcriptional regulation of gene expression *in vivo*. To better understand the mechanism of alternative polyadenylation that was observed for Id-2, CTE-1, MMP-9, Tyki and Fabp-4, it would be informative to perform ChIP studies on these genes in the attempt to correlate any changes in the phosphorylation state of RNAP-II CTD to the observed poly(A) site switch and to identify any changes in pre-mRNA processing factor association to the transcriptional machinery when the RAW cells are stimulated with LPS.

We have also identified AREs in the 3'-UTR of Id-2, CTE-1, MMP-9, Tyki and Fabp-4 that lie between the active promoter-proximal and promoter-distal poly(A) sites. Although there are no publications to date that identify these gene-specific elements as regulators of gene expression, all of them resemble regulatory elements previously identified to influence mRNA half-life or translation (150). Studies that elucidate the putative role these AREs play in gene expression on the mRNA transcript would provide important insight into the regulation of these genes and further support the importance of post-transcriptional gene regulation in the LPS-stimulated macrophage. Such studies would include the comparison of mRNA half-lives of the different mRNA isoforms of these genes under untreated and LPS-stimulated conditions. Furthermore, *in vivo* and *in vitro* protein binding studies to these putative AREs could elucidate

what mRNA binding proteins are responsible for the post-transcriptional regulation of these genes.

Since we have demonstrated that the 3'-end processing of pre-mRNA is a significant contributor to gene expression, altered regulation of this step can, therefore, disrupt cell homeostasis. For example, a single nucleotide mutation within the poly(A) site of the human prothrombin gene causes an increase in the pre-mRNA processing of that gene and subsequent over-accumulation of the thrombin protein causing thrombosis (52). In the case of malignancies of hematopoietic lineage, the t(11;14)(q13;32) translocation found most often in mantle cell lymphomas eliminates AREs in the 3'-UTR resulting in a profound increase in the mRNA stability and subsequent protein accumulation of the chimeric gene product (CCND1) resulting in neoplastic pathology (120). Even though the elimination of AREs in CCND1 does not involve alternative polyadenylation, cases may exist in other disease states where alternative polyadenylation does change the number of AREs within the 3'-UTR of a particular transcript. For example, the aberrant over-expression of the COX-2 enzyme is linked to many forms of malignant tumors. The 3'-UTR of the human COX-2 gene contains three functional polyadenylation sites (3) and multiple AREs that are highly conserved between human, mouse and rat (3, 41, 102). Multiple forms of the COX-2 message that correspond to alternative poly(A) site use have been detected in IL-1β treated human lung carcinoma cells and in the heart and lung of rats treated with LPS intraperitoneally (100). The proximal AREs have been shown to regulate translation while the distal AREs confer message stability (27). Interestingly, the incorporation of the distal ARE in the mature transcript is dependent on the poly(A) site chosen. Because of the nature of the COX-2 3'-UTR and the fact that alternative polyadenylation dictates the presence or absence of known mRNA instability elements in the mature transcript, it would

be noteworthy to identify the different mRNA isoforms and their half-lives present in various malignant tumors. Quite possibly, the increase in COX-2 protein expression can, in part, be attributed to an increase in message stability caused by alternative polyadenylation. Therefore, if various COX-2 mRNA isoforms are found in these tumors, studies on CstF-64 protein expression and 3'-end cleavage activity may be prudent to pursue. We could not assess the influence CstF-64 alone had on COX-2 gene expression because COX-2 is expressed only upon inflammatory or cytokine stimulation, and therefore was not detectable in unstimulated cells.

Interestingly, other studies from our lab have clearly shown that a poly(A) site switch does occur on the Ig heavy chain gene when B-cells differentiate into Ig-secreting plasma cells in the absence of CstF-64 increases (86). With the growing evidence that demonstrates transcription and pre-mRNA processing as linked events and that the splicing and 3'-end processing of gene transcripts is influenced by the presence of pre-mRNA processing factors on the migrating polymerase, we believe that CstF-64 is able to influence the selection of one poly(A) site over another through its local concentration to the nascent pre-mRNA. This could be achieved by increasing the over-all concentration of CstF-64 in the cell, which has been observed to be the limiting factor for active CstF trimer formation (86). However, we also believe that the local concentration of CstF-64 near the nascent pre-mRNA can be heightened through an increase in the recruitment of the steady state, constitutive levels of CstF-64 to the transcriptional machinery. Therefore, alternative polyadenylation can occur in conditions where no overt change in CstF-64 protein expression is observed through Western Blot of cellular lysates (86). As has been shown for other pre-mRNA processing factors in yeast, the recruitment of CstF-64 to the transcriptional machinery can be driven by a specific phosphorylation state of the RNAP-II CTD. Through our proposed studies to perform ChIP across the Ig heavy chain

locus in A20 and AxJ cells, we hypothesize that CstF-64 and/or other factors involved in pre-mRNA 3'-end processing are associated more with RNAP-II in the plasma cell line allowing for a greater selection of the weaker, promoter-proximal poly(A) site on the Ig transcript as it is expelled from the polymerase.

As we have clearly demonstrated in LPS-stimulated macrophages, the role of post-transcriptional gene regulation extends far beyond the popularly studied pro-inflammatory genes TNF-α and COX-2. We have shown that literally any gene, regardless of the number of poly(A) sites it possesses, can be influenced by the elevated levels of cleavage and polyadenylation brought on by increased CstF-64 expression. However, the more profound effects are demonstrated on genes that possess multiple poly(A) sites and AU-rich regulatory elements. We are also actively testing in B-lymphocytes the hypothesis that alternative polyadenylation can occur in the absence of increases in CstF-64 expression on the Ig heavy chain gene. Through the RNAP-II CTD phosphorylation-dependent recruitment of pre-mRNA processing factors (i.e. CstF-64) to the transcriptional machinery, we believe that alternative poly(A) site selection can occur on genes that possess multiple poly(A) sites without changing CstF-64 expression. This would extend the already existing dogma that links transcription to pre-mRNA processing to a greater understanding of the impact transcription and RNAP-II CTD phosphorylation bestow on post-transcriptional gene regulation.

Table 1. Sequences of oligonucleotides used for semi-quantitative RT-PCR and real-time QPCR

Gene Name		Sequence (5'-3')
TNF-α	Sense	CAC GTC GTA GCA AAC CAC CAA GTG GA[a]
	Antisense	TGG GAG TAG ACA AGG TAC AAC CC
iNOS	Sense	CAC CTT GGA GTT CAC CCA GT
	Antisense	ACC ACT CGT ACT TGG GAT GC
COX-2	Sense	AAA CCG TGG GGA ATG TAT GAG C[a]
	Antisense	TCG CAG GAA GGG GAT GTT GTT C
CHOP	Sense	GCA TGA AGG AGA AGG AGC AG
	Antisense	CTT CCG GAG AGA CAG ACA GG
Id-2	Sense	ATC AGC CAT TTC ACC AGG AG
	Antisense	TCC CCA TGG TGG GAA TAG TA
Tyki	Sense	AGC TCT GCA ATC CCG TTC TA
	Antisense	CCA GAA TCC ACG GTG CTT AT
MMP-9	Sense	TGA ATC AGC TGG CTT TTG TG
	Antisense	GTG GAT AGC TCG GTG GTG TT
PLOD-2	Sense	GTC AAG CAG AAG CCA GGA AC
	Antisense	CCA CAA CTT TCC ATG ACG TG
P4ha2	Sense	AGC GAT GAC GAA GAT GCT TT
	Antisense	CCA CTT GTT GGA GAC CCA CT
MMP-12	Sense	TTT CTT CCA TAT GGC CAA GC
	Antisense	GGT CAA AGA CAG CTG CAT CA
RARβ	Sense	ACA AGT CAT CGG GCT ACC AC
	Antisense	CTG TGC ATT CCT GCT TTG AA
CTE-1	Sense	GAT GGC CTC AAG GAT GTT GT
	Antisense	TAA GGG GGC TCG ATG TAA TG
CD48	Sense	CGA TCT CAT CGT CAC ACC AC
	Antisense	TGC TTC CAA GTT GCC TTC TT
HPRT	Sense	CGT CGT GAT TAG CGA TGA TG
	Antisense	AGA GGT CCT TTT CAC CAG CA

[a] Primer sequences obtained from (56).

Table 2. Genes that exhibit increased expression in CstF-64 over-expressing RAW macrophages[a]

Accession Number	Gene Name	Fold Increase[b]		No. of Putative Poly(A) Sites[c]
		CstF-64	LPS	
Biological Regulation/Metabolism				
NM_020557	Tyki	2.64	35.12	3
NM_013599	MMP-9	2.05	7.11	2
NM_011961	PLOD-2	17.62	-	1
BC018411	P4ha2	13.01	-	1
BC019135	MMP-12	5.97	-	1
NM_080555	Ppap2b	4.88	-	ND[d]
Transcriptional Regulation				
NM_007837	CHOP	2.81	7.22	ND
NM_010496	Id-2	2.64	3.45	4
NM_011243	RARβ	4.93	-	2
Transport				
BC002148	Fabp-4	4.76	5.11	2
NM_023044	Slc15a3	2.40	20.09	ND
NM_007752	Ceruloplasmin (Cp)	2.40	17.98	1
NM_007443	Bikunin	5.29	-	1
Signaling				
AK089198	GPCR-35	4.90	-	5
NM_022881	RGS-18	3.59	-	2
Development				
NM_019733	Hermes	14.81	3.22	1
AF220524	DNMT3L	2.84	-	ND
Other Genes/Unknown				
AY239223	Epsti1	2.78	5.44	1
BC022751	IFN-stimulated protein	2.29	22.78	ND
BC057864	Digr2	2.15	7.31	1
BC010564	histone H2a	2.09	7.83	ND
BC024822	EST AW146242	3.79	-	2
BC009118	cDNA sequence BC009118	2.51	-	1

[a] Data are from three independent experiments and were analyzed as described in *Materials and Methods*. Genes shown were induced an average of 2-fold or greater.
[b] Fold increase in average signal value observed in CstF-64 over-expressing or LPS-stimulated RAW macrophages compared to untreated, control.
[c] Number of AAUAAA and AUUAAA poly(A) signals with associated GU-rich regions identified in the 3'-UTR of the listed GenBank sequence.
[d] ND – no identifiable AAUAAA or AUUAAA poly(A) site in the 3'-UTR of the listed GenBank sequence.

Table 3. Genes that exhibit decreased expression in CstF-64 over-expressing RAW macrophages[a]

Accession Number	Gene Name	Fold Decrease[b]		No. of Putative Poly(A) Sites[c]	Accession Number	Gene Name	Fold Decrease[b]		No. of Putative Poly(A) Sites[c]
		CstF-64	LPS				CstF-64	LPS	
Biological Regulation/Metabolism					**Transport**				
AK053680	CTE-1	6.71	3.13	4	NM_007643	CD36	16.16	-	ND
NM_134188	MTE-1	5.19	2.27	1	AF057367	Uty	5.92	-	4
NM_012011	eIF-2α Y-linked	14.74	-	5	**Signaling**				
NM_020026	B3galt3	7.59	-	1	BC010200	FGFR1	4.45	-	1
NM_153153	Supervillin	4.08	-	2	NM_010336	Edg2	3.81	-	2
NM_007945	Eps8	3.04	-	2	AB122023	Pilr-1	3.37	-	1
AK048950	Siat10	2.95	-	3	AK028875	CXCL16	2.71	-	3
NM_011890	sarcoglycan, beta	2.10	-	4	NM_016846	ralGDSL-1	2.67	-	2
BC048929	Dystonin	2.01	-	3	U66888	EMR-1	2.29	-	1
Adhesion/Binding					NM_152915	DNER	2.12	-	3
NM_021495	Nectin 3	2.21	2.38	ND[d]	NM_008967	Ptgir	2.11	-	1
NM_012008	Ddx3 Y-linked	16.06	-	7	**Development**				
NM_011419	Jumanji 1D	5.48	-	3	BC016887	DOC-2	2.42	-	4
NM_029413	Zcwcc2	4.19	-	ND	**Other Genes**				
NM_007965	Evl	4.16	-	1	NM_153098	CD109	5.40	-	2
AF116823	Trip6	3.39	-	ND	BC060977	CD48	4.30	-	2
AK030971	coronin 2A	2.03	-	1					

[a] Data are from three independent experiments and were analyzed as described in *Materials and Methods*. Genes shown were induced an average of 2-fold or greater.
[b] Fold increase in average signal value observed in CstF-64 over-expressing or LPS-stimulated RAW macrophages compared to untreated, control.
[c] Number of AAUAAA and AUUAAA poly(A) signals with associated GU-rich regions identified in the 3'-UTR of the listed GenBank sequence.
[d] ND – no identifiable AAUAAA or AUUAAA poly(A) site in the 3'-UTR of the listed GenBank sequence.

Table 4. Sequences of oligonucleotides used for MRP-PCR

Gene Name	Primer Name	Primer Sequence	Product Size with F
Id-2	F	CCA GCA TCC CCC AGA ACA AG	
	R1	GGG AAT TCA GAT GCC TGC AA	211
	R2	GCT TGC TTT ATT TCA GAC AAC CAG TG	937
CTE-1	F	CTG GTG GGT GCC AAC ATC AC	
	R1	CCC AGC CCT TGA ATC AGC AC	290
	R2	CGG GAA TGA GAG CGT TTG CT	937
MMP-9	F	CTG GAC TCC GCC TTT GAG GA	
	R1	GAA AGG ACG GTG GGC AGA GA	623
	R2	CAG CTG CCA GCA ACC ACA GT	1204
Tyki	F	TGC CCT GAG TGA GTG CCT TG	
	R1	CAG AGC CAC CAG GCT TCT CC	477
	R2	TTC AGC CTC TGG TGC AAT CA	1163
Fabp-4	F	ATT ACC GTG CCC ACC TGC AT	
	R1	GCC CCG CCA TCT AGG GTT AT	320
	R2	TAT GGG AGG GTG GAG CCT GA	676

Table 5. Groups of AU-rich Elements (AREs)[a]

Group	Motif	Example
I	WAUUUAW and a U-rich region	c-fos, c-myc
IIA	AUUUAUUUAUUUAUUUAUUUA	GM-CSF, TNF-a
IIB	AUUUAUUUAUUUAUUUA	Interferon-a
IIC	WAUUUAUUUAUUUAW	cox-2, IL-2, VEGF
IID	WWAUUUAUUUAWW	FGF2
IIE	WWWWAUUUAWWWW	u-PA receptor
III	U-rich, non-AUUUA	c-jun

[a] adapted from (150).

Table 6. Sequences of Oligonucleotides Used For 5'-RACE PCR and ChIP

5'-RACE PCR		Sequence (5'-3')
GSP1	Antisense	GCC AGT GAG GGA CAA AGA AAG CAT AG
GSP2	Antisense	CCA ATA GTC GCA GGC GGA ATA ATC AC
IgG2a Locus[a]		
Promoter	Sense	ATG CAA ATC CTG TGG GTC TAC A
	Antisense	CAG CGA TCT CAG GGT TTC TGT
V-region	Sense	CGG CCA CAT TAA CTG TTG ACA
	Antisense	TCG CAG GCG GAA TAA TCA C
IgH Eμ	Sense	GGC CGA TCA GAA CCA GAA CA
	Antisense	CCA AAT AGC CTT GCC ACA TGA
3'-IgH Eμ	Sense	CAC AGT AAT GAC CCA GAC AGA GAA G
	Antisense	CCT CGC TTA CTA GGG CTC TCA A
IgH C1	Sense	CCC TGT GTG TGG AGA TAC AAC TG
	Antisense	TCA CTG GCT CAG GGA AAT AAC C
IgH M1	Sense	CCT GCT CAG CGT GTG CTA CA
	Antisense	CAG CAG AGG GTG GGA GAC A
3'-pAm	Sense	CCA GAA GAG AAG CGT GTG GTA CT
	Antisense	GGA TCC ATG CTC TTT GCT GAA
Control Gene[a]		
MyoD1 promoter	Sense	AGA GTC CAG GCC AGG GAA GA
	Antisense	TGT GGA GAT GCG CTC CAC TA
GAPDH promoter	Sense	GAG ACA GCC GCA TCT TCT TGT
	Antisense	CAC ACC GAC CTT CAC CAT TTT
β2-Globin promoter	Sense	CCT GCT AGA AGC AAG GTC AGA AA
	Antisense	TTT CAG GAT TCC GTG TGT AAA AG
IgG2a Promoter Sequencing		
M13	Sense	CAG GAA ACA GCT ATG AC
	Antisense	GTA AAA CGA CGG CCA G

[a] All cDNA sequences used for primer design using Primer Express were obtained from the GenBank database.

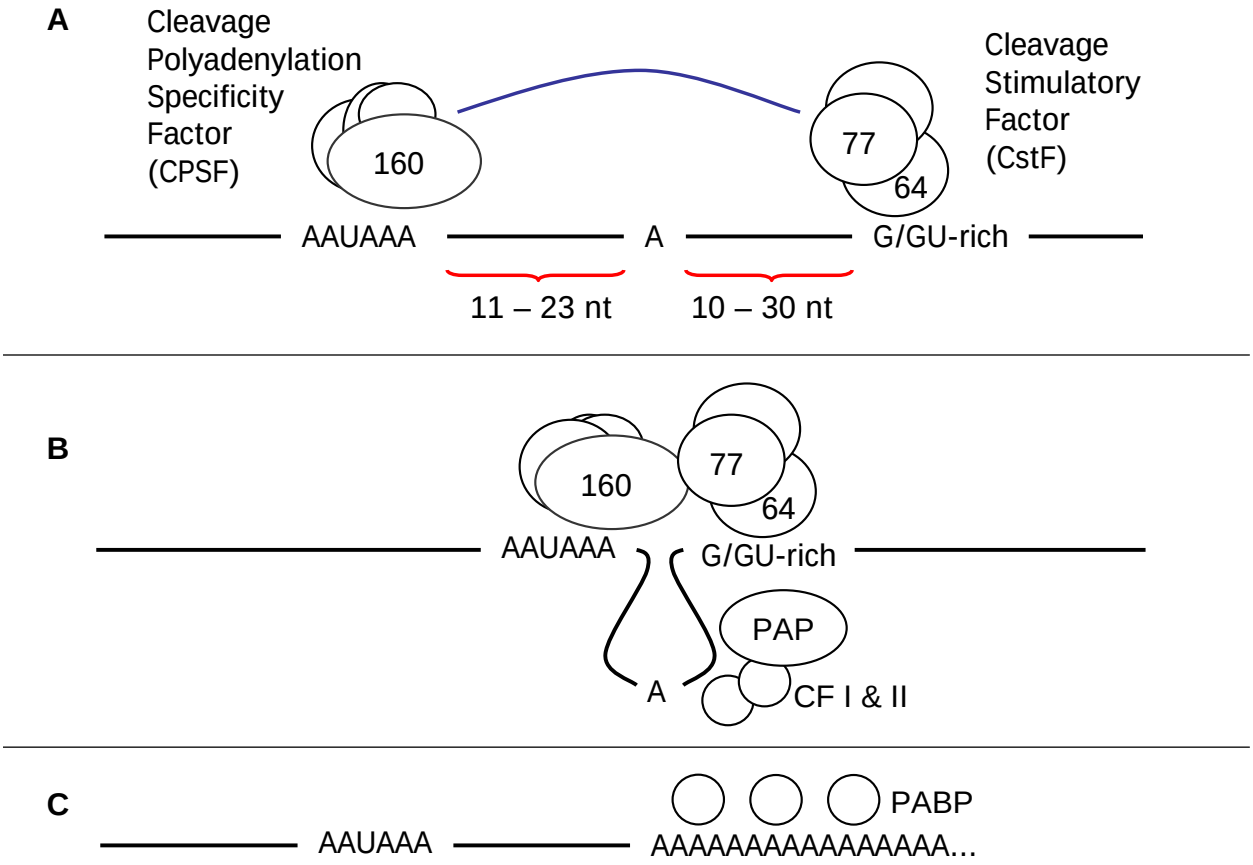


Figure 1. Mechanism of pre-mRNA 3'-end Cleavage and Polyadenylation

A. Cleavage Polyadenylation Specificity Factor (CPSF) binds to the consensus AAUAAA poly(A) signal through it's 160-kDa subunit, CPSF-160. Cleavage Stimulatory Factor (CstF) binds to the less conserved GU-rich region downstream from the cleavage site through it's 64-kDa subunit, CstF-64. Studies have shown that stronger polyadenylation sites have a poly(A) signal approximately 11-23 nucleotides upstream of the cleavage site and the GU-rich region approximately 10-30 nucleotides downstream from the cleavage site. B. Protein/protein interactions between CPSF-160 and the 77 kDa subunit of CstF, CstF-77, cause an out-looping of the intervening RNA that exposes an adenosine residue to cleavage and subsequent polyadenylation. *In vitro* studies have shown that Cleavage Factors I and II (CF I and CF II) and Poly(A) Polymerase (PAP) are also required to complete the pre-mRNA 3'-end processing. C. Coincident with the addition of the first adenosine residues of the poly(A) tail, Poly(A) Binding Proteins (PABP) assemble onto the poly(A) tail and confer mRNA stability and induce rapid, processive addition of the remaining adenosines.

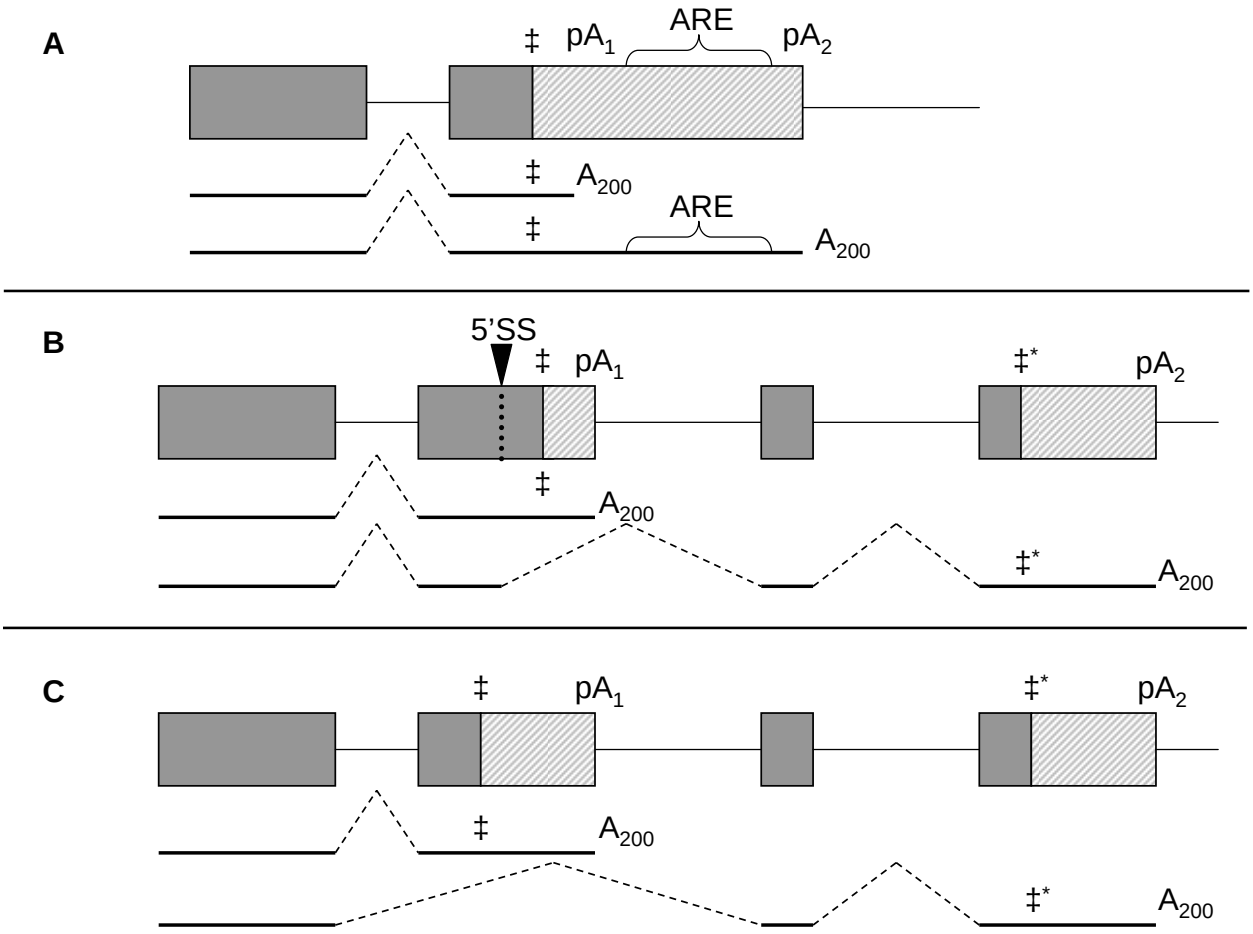


Figure 2. Comparison of Three Motifs of Exon Arrangement in the Context of Alternative Polyadenylation.

Solid dark gray boxes represent exons and the thin line(s) that connect them represent introns. Light gray, hatched boxes indicated the 3'-UTR upstream of the indicated poly(A) site. Below each exon arrangement are depicted the two possible mRNA transcripts stemming from the particular splicing pattern shown (indicated by the dashed lines) and the use of either the promoter-proximal poly(A) site (pA_1) or promoter-distal poly(A) site (pA_2). A. Tandem poly(A) sites. In this exon arrangement, all alternative poly(A) sites are located downstream from the 3'-most protein coding exon (‡). Thus, selection of one poly(A) site over another is disconnected from splicing mechanisms and determines whether AU-rich regulatory elements (ARE) are retained in the mature transcript. B. Composite internal/terminal exons. In this exon arrangement, there exists a competition between the splicing and polyadenylation reactions that results in an mRNA that codes for proteins possessing different carboxy-termini. When the cleavage and polyadenylation reaction at the proximal poly(A) site (pA_1) is chosen before or instead of the splicing reaction at the 5'-internal splice site within the preceding exon (5'SS), then the short transcript is formed that possesses the complete terminal exon (‡). If the 5'SS is selected before cleavage/polyadenylation can occur at pA_1, then a composite exon is created and a different terminal exon is incorporated into the final transcript (‡*). Polyadenylation subsequently occurs at the distal poly(A) site (pA_2). C. Skipped Exons. The final mRNA product for this exon arrangement is dictated primarily by the splicing pattern of the transcript. Cleavage and polyadenylation will occur at pA_1 only if the exon immediately preceding the poly(A) site is incorporated into the final gene product. Because this exon does not possess a splice site donor, no further splicing will occur and will thus be the terminal exon of the transcript (‡). However, if the exon that is associated with pA_1 is not chosen for incorporation into the final gene product (i.e. it is skipped), then the distal poly(A) (pA_2) site will be chosen for 3'-end processing and a different terminal exon will be in the final transcript (‡*). Figure adapted from (45)

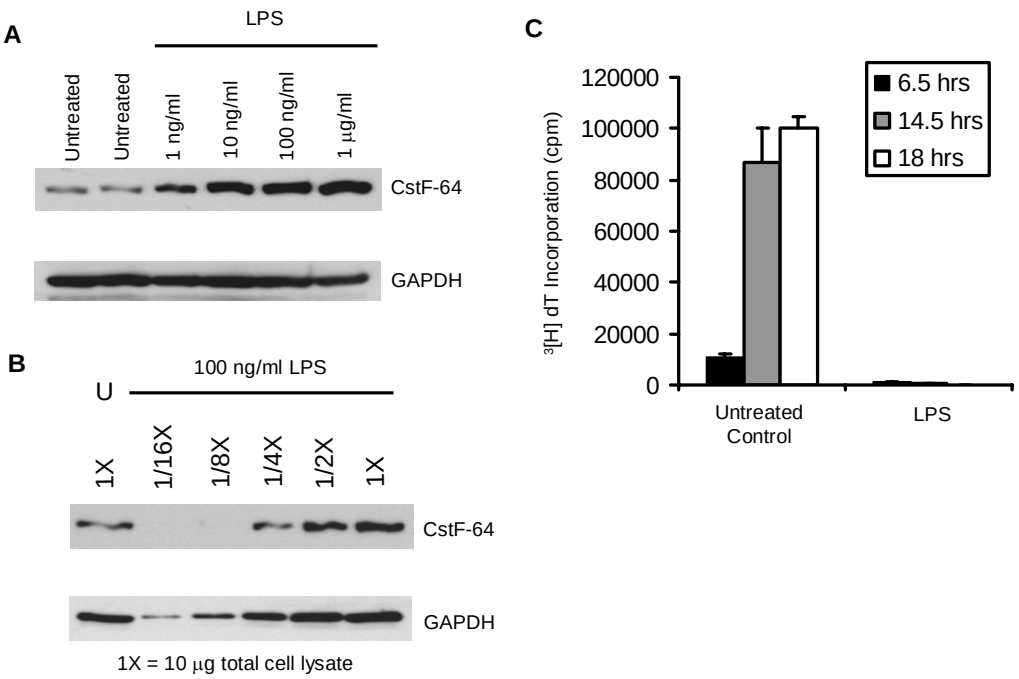


Figure 3. LPS stimulation of RAW 264.7 macrophages increases CstF-64 protein expression.

A. Total protein was harvested from macrophages stimulated for 18 hours with increasing amounts of LPS. CstF-64 protein expression was assayed by Western blot. B. A dilution series of the protein extracts from the macrophage cell line stimulated for 18 hours with 100 ng/ml LPS was performed for semi-quantitative determination of LPS-induced CstF-64 induction by Western blot (1x = 10 μg, 1/2x = 5 μg, 1/4x = 2.5 μg, 1/8x = 1.25 μg, and 1/16x = 0.625 μg). The diluted LPS-stimulated samples were compared to 10 μg of protein extract from untreated cells. C. Macrophages were cultured in the absence or presence of 100 ng/ml LPS and pulsed with ^{3}[H] thymidine at various time points to determine cellular proliferation. Averages ± SEM are shown for sample size n=3.

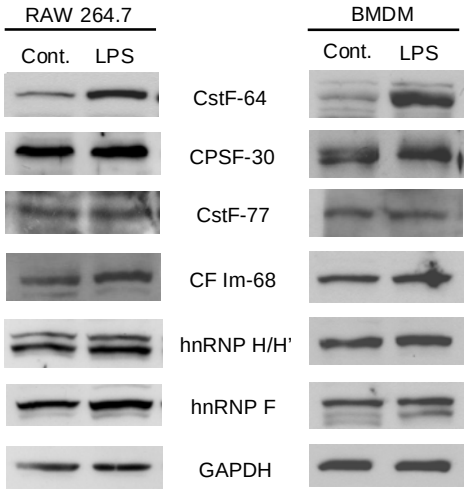


Figure 4 LPS stimulation of RAW 264.7 macrophages and BMDM does not alter the expression levels of other factors involved in the cleavage/polyadenylation reaction.

Protein extracts from macrophages stimulated for 18 hours with 100 ng/ml of LPS were probed by Western blot for the indicated cleavage/polyadenylation factors.

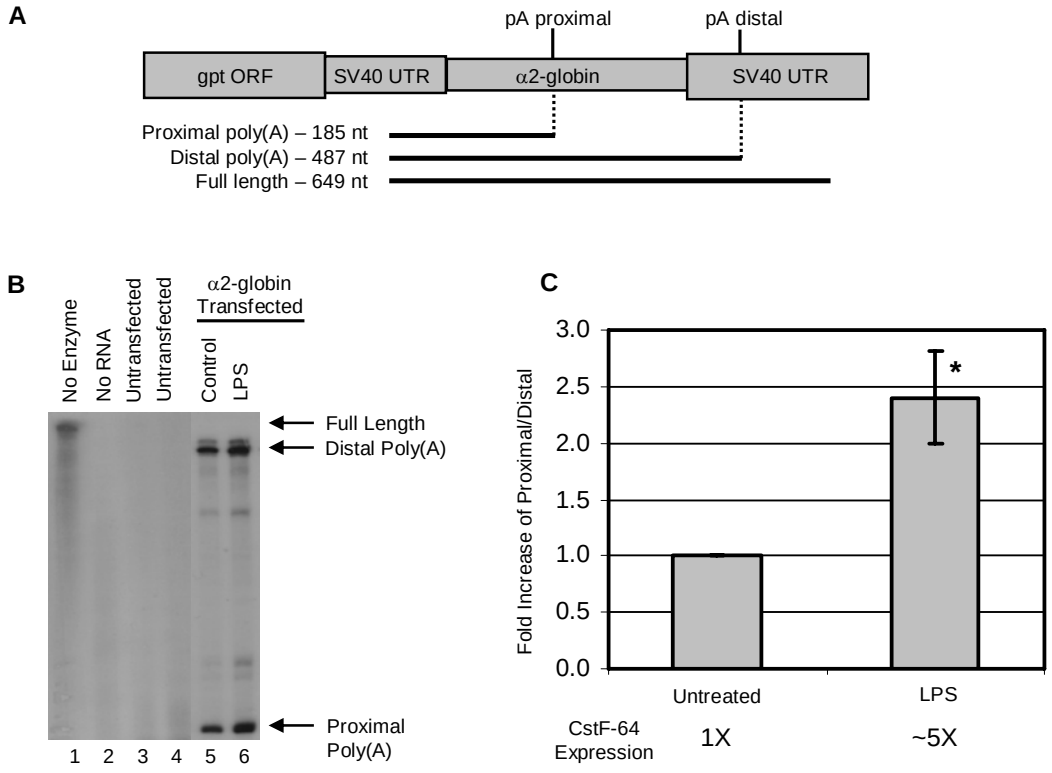


Figure 5. Stimulation of RAW 264.7 macrophages with LPS caused increased selection of a weaker promoter-proximal poly(A) site.

A. Schematic diagram of the α2-globin reporter construct. Predicted sizes of the protected antisense riboprobe fragments, based on the poly(A) site chosen on the reporter construct, are indicated. B. mRNA from macrophages stimulated with LPS was used in an RNase protection assay to measure changes in poly(A) site choice. Controls include probe without T2 RNase to show the full length probe (lane 1), probe without RNA plus T2 RNase to illustrate complete digestion of the probe in the absence of RNA (lane 2), and probe with RNA from untransfected macrophages to control for nonspecific hybridization (lanes 3 and 4). Lanes 5 and 6 show protected riboprobe fragments from RAW-α2 RNA unstimulated or LPS-stimulated, respectively. C. Fold changes in P/D ratios of macrophages stimulated with LPS are compared to the P/D ratio achieved from untreated macrophages. Below the graph indicates the relative amount of CstF-64 protein measured from the same cultures the RNA was harvested from. In using the paired student-t test, significant changes in poly(A) site selection were detected in LPS-stimulated macrophages. Averages ± SEM are shown based on sample size n=5. * $p<0.005$.

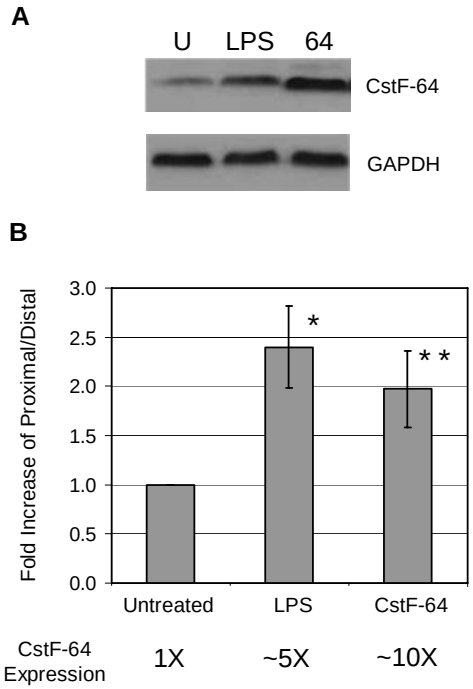


Figure 6. Over-expression of CstF-64 in RAW 264.7 macrophages caused increased selection of a weaker promoter-proximal poly(A) site.

A. Western blot comparing the levels of CstF-64 between untreated (U), LPS-treated (LPS), and CstF-64 over-expressing (64) RAW macrophages. B. Fold changes in P/D ratios of macrophages either stimulated with LPS or over-expressing CstF-64. Below the graph indicates the relative amount of CstF-64 protein measured from the same cultures the RNA was harvested from. In using the paired student-t test, significant changes in poly(A) site selection were detected in LPS-stimulated and CstF-64 over-expressing macrophages. Averages ± SEM are shown based on sample size n=5. * $p<0.005$. ** $p<0.01$.

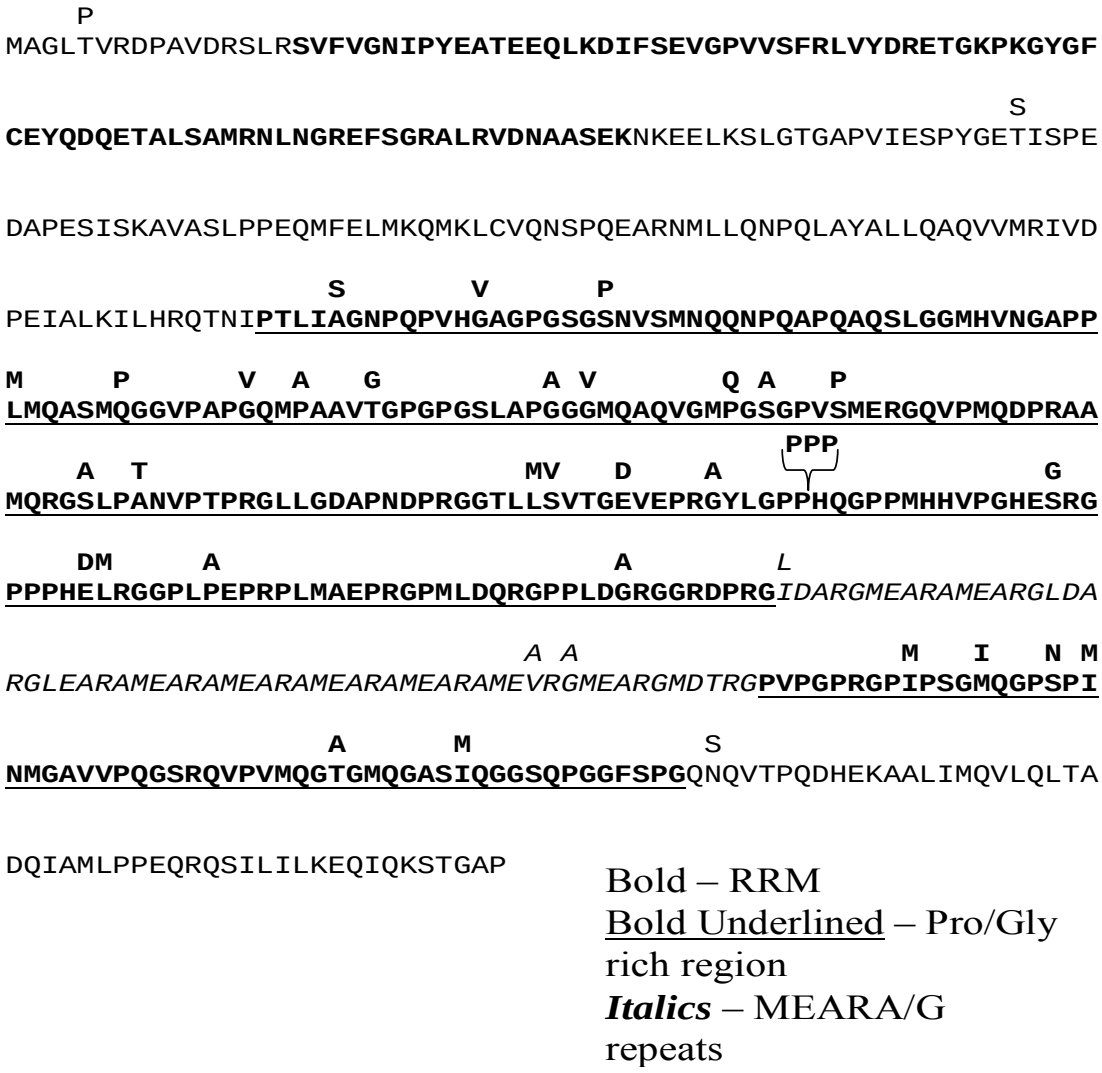


Figure 7. Comparison of Human and Mouse CstF-64 Amino Acid Sequence

Shown is the amino acid sequence for the human form of CstF-64 (NM_001325) that we over-expressed in the RAWα2 macrophages. Letters that appear above the primary sequence indicate the amino acid change that appears in the mouse form of the protein for that residue. The primary sequence is font coded to identify structural and functional motifs of the protein. Bold residues identify the conserved RNA recognition motif (RRM) shared by CstF-64 and many other RNA binding proteins. The first block of normal font residues identifies the hinge region that is responsible for binding to CstF-77. Emboldened and underlined residues identify proline/glycine-rich region, while the italics residues identify the conserved MEARA/G repeats responsible for the alpha-helical portion of the protein. The last block of normal font residues identifies the C-terminal domain of CstF-64, a domain that exhibits affinity to the PC4 transcription elongation factor. A triple proline insert appears in the mouse form and is located immediately downstream of Proline 350, as shown.

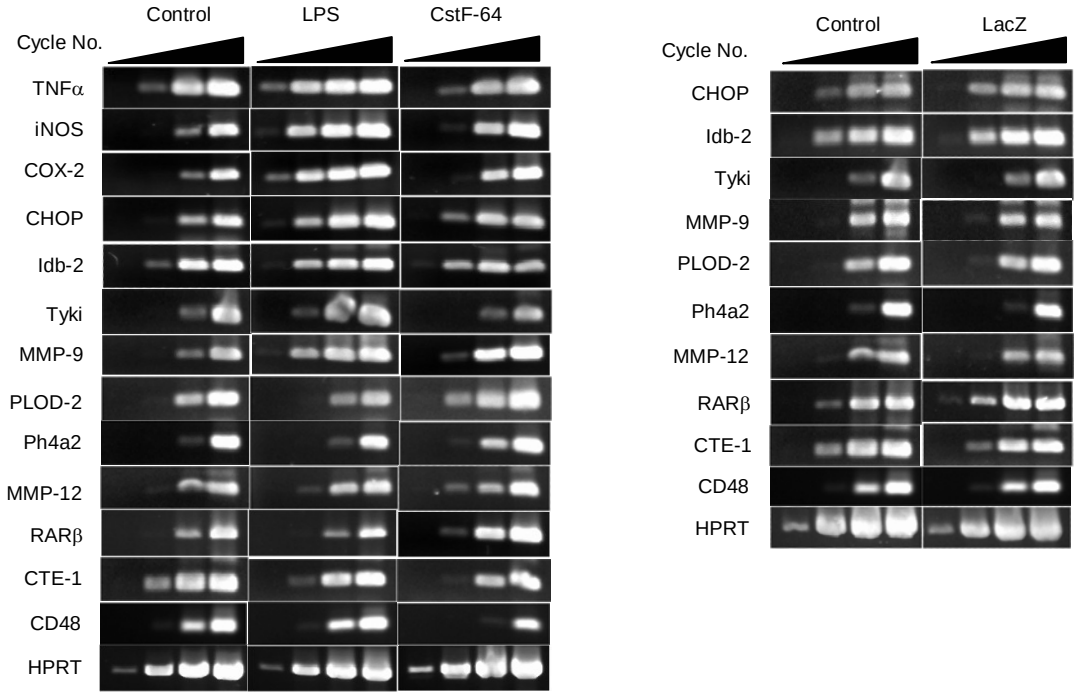


Figure 8. Validation of differential gene expression observed in LPS-stimulated and CstF-64 over-expressing RAW 264.7 macrophages by semi-quantitative RT-PCR

Ethidium bromide-stained agarose gels show increases in product formation from 20, 25, 30 or 35 PCR cycles which ensures that the reactions lie within the linear range, and therefore can be accurately compared between RAW macrophage culture conditions.

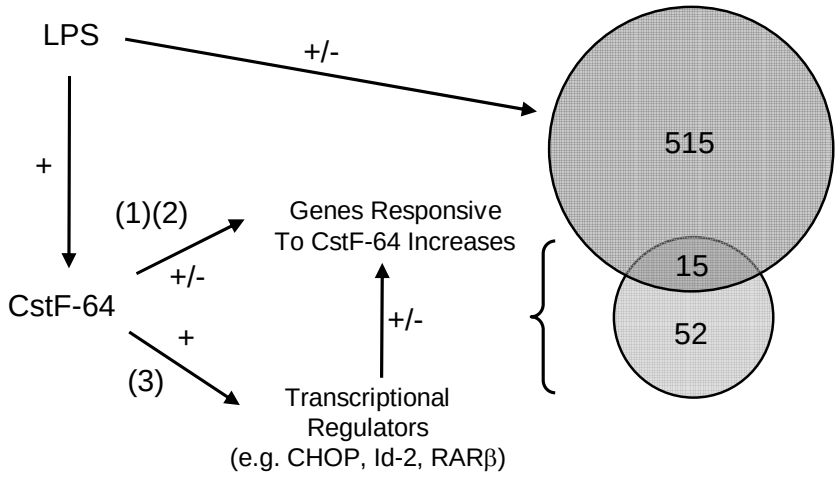


Figure 9. Model of CstF-64 influences on gene expression.

LPS treatment of RAW macrophages increases CstF-64 protein levels. Depicted are three possible ways elevated levels of CstF-64 protein influence gene expression: (1) The cleavage activity of the cell is increased allowing a larger number of pre-mRNA transcripts to be 3'-end cleaved and polyadenylated, especially for transcripts that have GU-rich sequences in the 3'-UTR that are sub-optimal for CstF-64 binding. This would create more mature mRNA transcripts of a given gene to be used for protein translation. (2) An alternative polyadenylation event occurs on a gene transcript that has multiple poly(A) sites surrounding mRNA elements that control stability. Increased selection of one poly(A) site over another would determine the presence or absence of these mRNA stability elements from the mature transcript, thus increasing mRNA half-life and allowing the accumulation of more protein product. (3) The increased expression of transcriptional regulators (e.g. CHOP, Id-2, RARβ) by either of the two aforementioned mechanisms has effects on the expression on their targeted genes. The Venn diagram depicts 515 genes that change expression with LPS stimulation, 52 genes that change expression with CstF-64 over-expression, and 15 genes that change under both conditions. "+" Induction of gene expression; "–" Repression of gene expression.

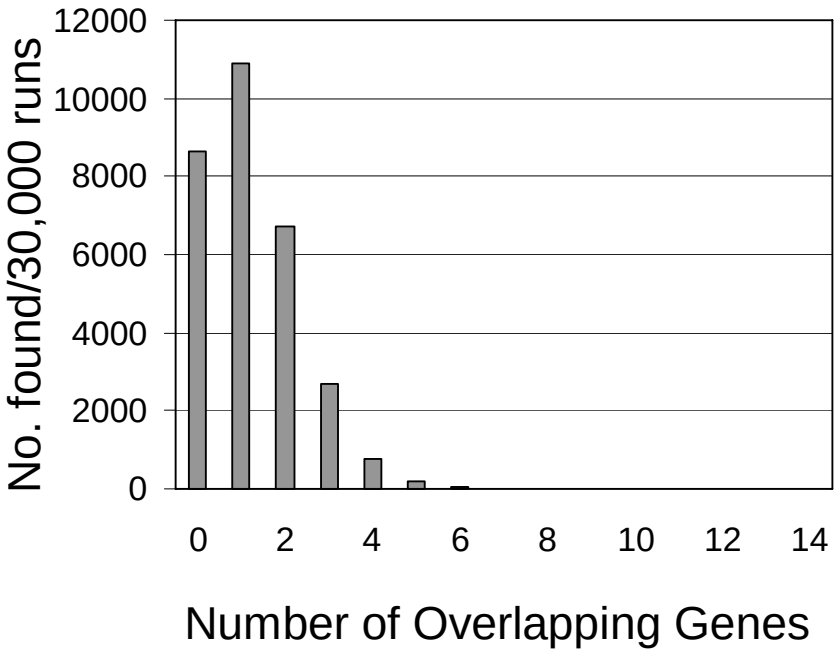


Figure 10. The number of overlap genes detected by microarray analysis does not occur by random chance.

A Statistical algorithm was designed to test the random chance of the same 15 genes out of 22,690 on the array being selected from two separate pools at random. We found that this overlap between the LPS stimulation and CstF-64 over-expression would occur far less than once in 30,000 trials.

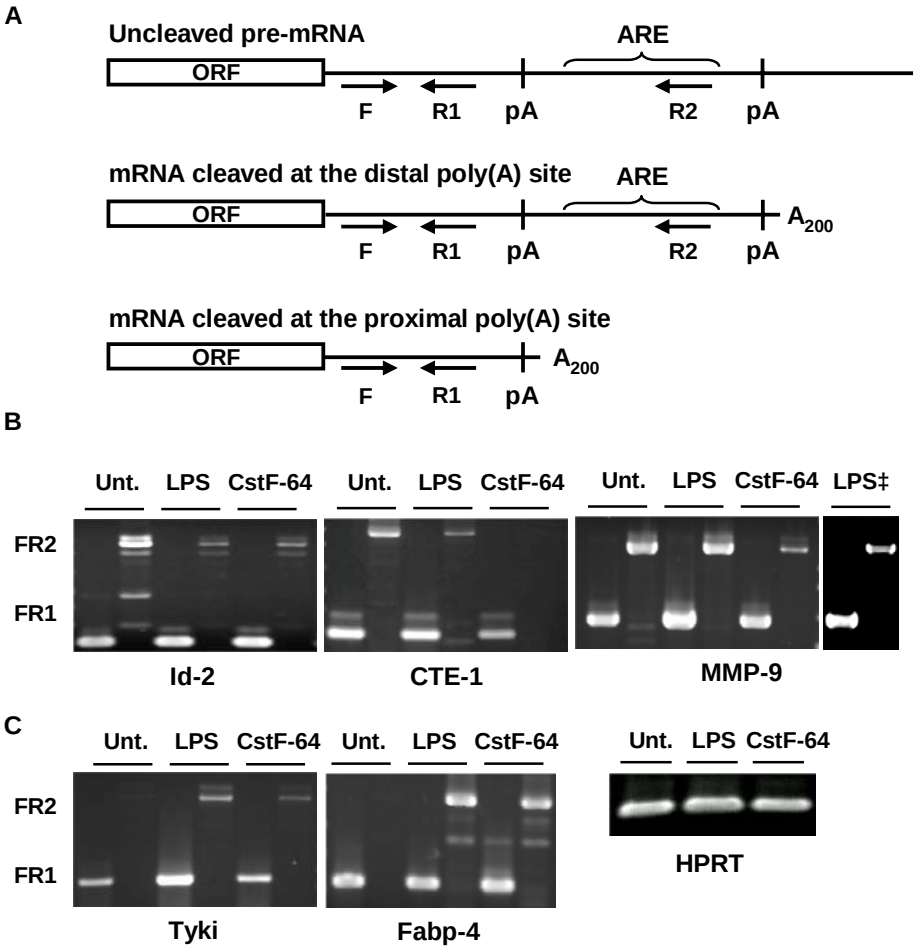
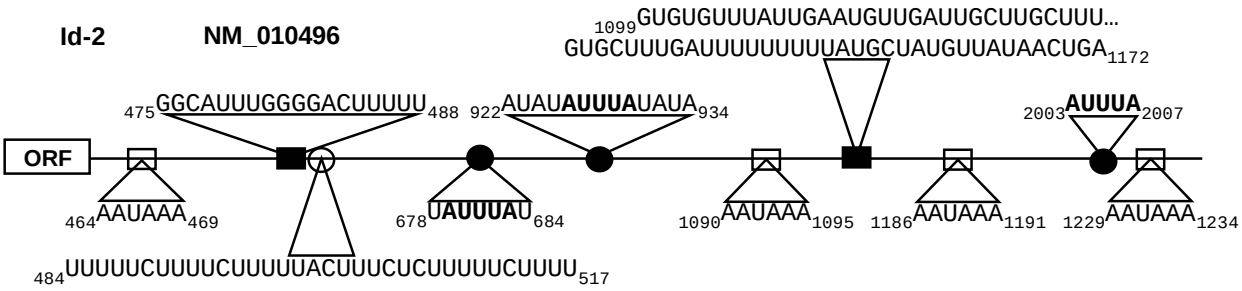


Figure 11. Increased levels of CstF-64 protein induce poly(A) site switch.

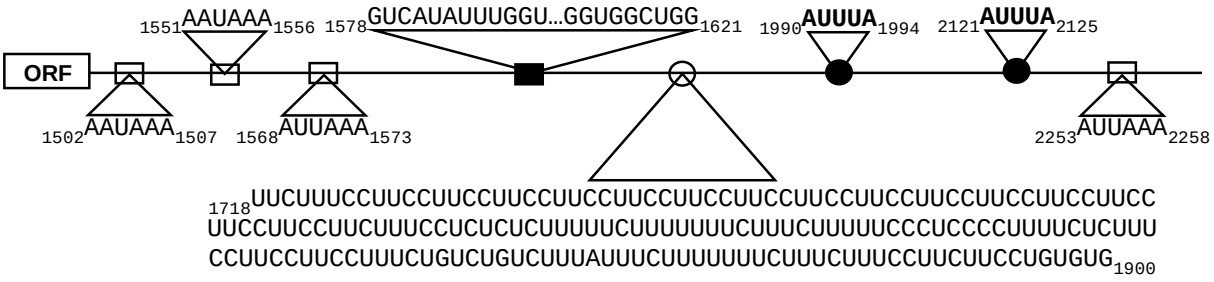
A. Schematic diagram of multiple-reverse primer PCR (MRP-PCR) experimental approach. The diagram on the top depicts the locations of the common forward primer (F), the common reverse primer (R1) and the reverse primer that is specific for transcripts only processed at the putative distal poly(A) site (R2). Also shown are the locations of the two putative poly(A) sites and the AU-rich elements (ARE) that lie between them. The middle and bottom diagram depict the mature gene product when it is processed at the distal or proximal poly(A) site, respectively. PCR using the F and R1 primers (denoted FR1) will amplify all gene products regardless of where they are processed. PCR using the F and R2 primers (denoted FR2) will amplify only the gene products that are processed at the downstream poly(A) site. B. Ethidium bromide-stained agarose gels show decreases in distal poly(A) site use under LPS-stimulating and CstF-64 over-expressing conditions for Id-2, CTE-1 and MMP-9. ‡. A lighter exposure of the gel for MMP-9 under LPS-stimulating conditions is also shown to better compare the reduced amount of FR2 PCR product produced relative to FR1. C. Ethidium bromide-stained agarose gels show increases in distal poly(A) site usage under LPS-stimulating and CstF-64 over-expressing conditions for Tyki and Fabp-4. Also shown is an agarose gel of HPRT that demonstrates that equal amounts of cDNA were applied to all PCR reactions.

Gene Name **Accession No.**

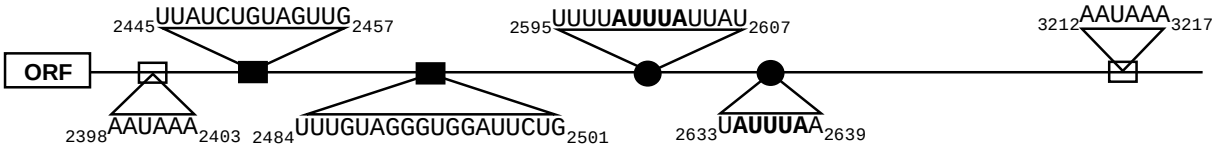
Id-2 **NM_010496**



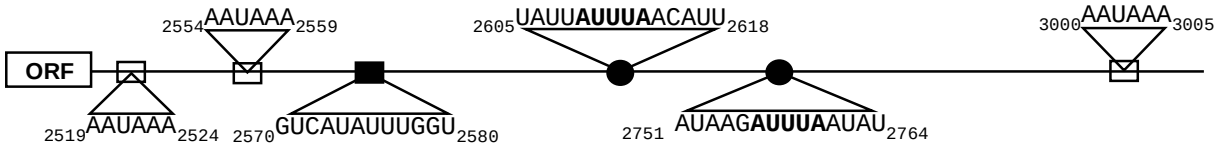
CTE-1 **AK053680**



MMP-9 **NM_013599**



Tyki **NM_020557**



Fabp-4 **BC002148**

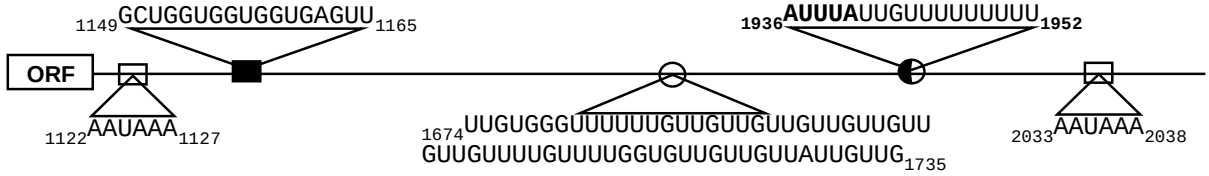


Figure 12 Diagram depicting key elements of the Id-2, CTE-1, MMP-9, Tyki and Fabp-4 3'-UTRs.

ORF indicates the 3'-end of the transcript's open reading frame. Precise locations of the elements are indicated and derived from the Entrez accession number provided. Open boxes indicate the consensus AAUAAA or nonconsensus AUUAAA poly(A) signal and where CPSF-160 binds. Closed boxes indicate the GU-rich region that lies downstream from the poly(A) signal and where CstF-64 binds. The open circle indicates a non-AUUUA containing U-rich region. Closed circles indicate the putative AU-rich elements. Emboldened within these elements is the core AUUUA consensus sequence. The half-filled circle in Fabp-4 3'-UTR shows a core AUUUA consensus sequence with an associated U-rich stretch. For the Id-2, CTE-1 and MMP-9 genes, increased use of the proximal poly(A) site was observed under LPS-stimulating and CstF-64 over-expressing conditions. As a result, the putative AREs indicated would be removed from the final gene product, thus eliminating any regulatory influences these elements may have on the half-life or translatability of the mRNA. For both the Tyki and Fabp-4 genes, increased use of the distal poly(A) site was observed under LPS-stimulating and CstF-64 over-expressing conditions. As a result, the putative AREs indicated would be added to the final gene product, thus incorporating any regulatory influences these elements may have on the half-life or translatability of the mRNA.

NM_133196

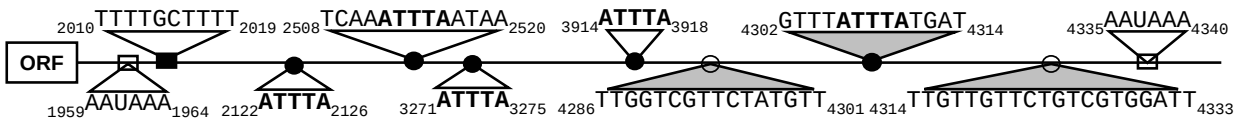


Figure 13. Mouse CstF-64 mRNA 3'-Untranslated region

ORF indicates the 3'-end of the transcript's open reading frame. Precise locations of the elements are indicated and derived from the Entrez accession number provided. Open boxes indicate the consensus AAUAAA poly(A) signal and where CPSF-160 binds. Closed boxes indicate the GU-rich region that lies downstream from the poly(A) signal and where CstF-64 binds. Closed circles indicate the putative AU-rich elements. Emboldened within these elements is the core AUUUA consensus sequence. The open circle indicates a non-AUUUA containing U-rich region. The 3'-UTR of the mouse CstF-64 gene transcript possesses a variety of putative AREs. The AREs with the shaded triangles that lie immediately upstream of the distal poly(A) site are of most interest when determining the potential translation regulatory role they bestow on CstF-64 gene expression, although any of the putative AREs may contribute to the translatability of the gene transcript.

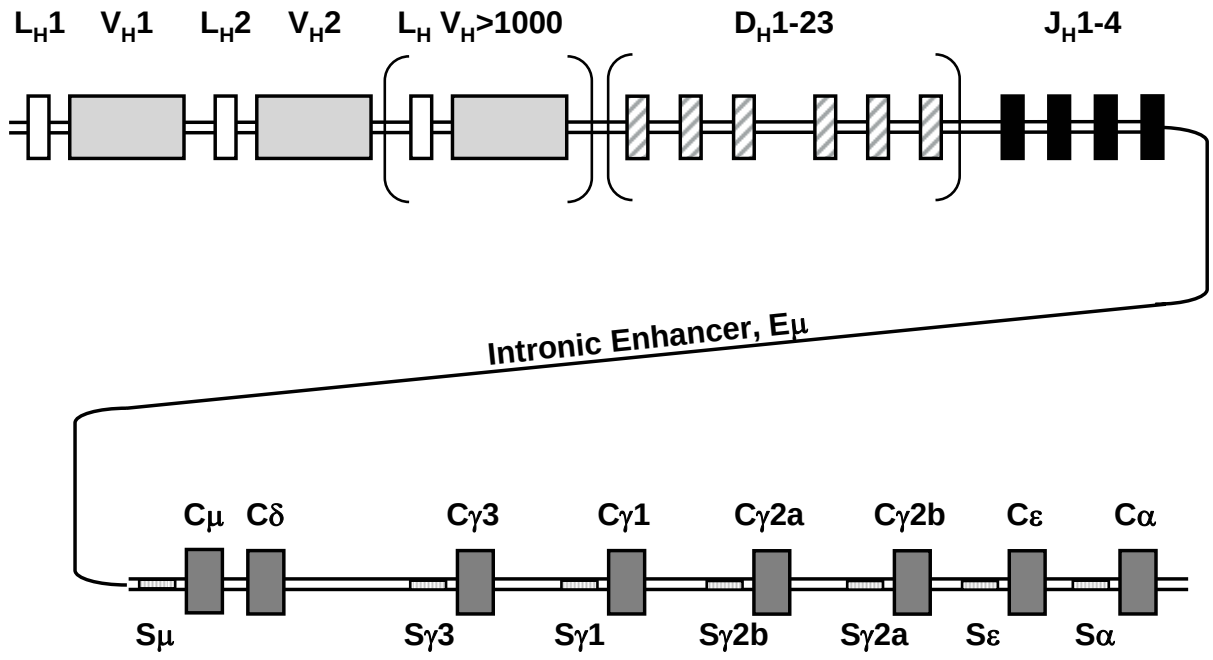


Figure 14. Mouse immunoglobulin germline genomic arrangement

The germline genomic arrangement of the mouse immunoglobulin gene is indicated prior to VDJ recombination and isotype switching. There are more than 1000 different heavy chain variable regions (VH) each associated with its own leader sequence (L_H) and promoter (not shown). Downstream of the 3'-most V_H lies the diversity region (D_H) cluster that contains 23 alternative segments followed by the joint region (J_H) cluster that contains 4 segments. Located approximately 500 nucleotides downstream from the 3'-most J_H is the intronic enhancer (Eμ) that contributes to B-cell development and Ig transcription. Following Eμ is the heavy chain constant region locus beginning with C_μ and followed by C_δ, $C_{\gamma 1}$, $C_{\gamma 2a}$, $C_{\gamma 2b}$, C_ε and C_α, respectively (not shown to scale). Each constant region locus possesses its own switch region (S) located immediately upstream of the constant region locus, except C_δ which shares the switch region with C_μ.

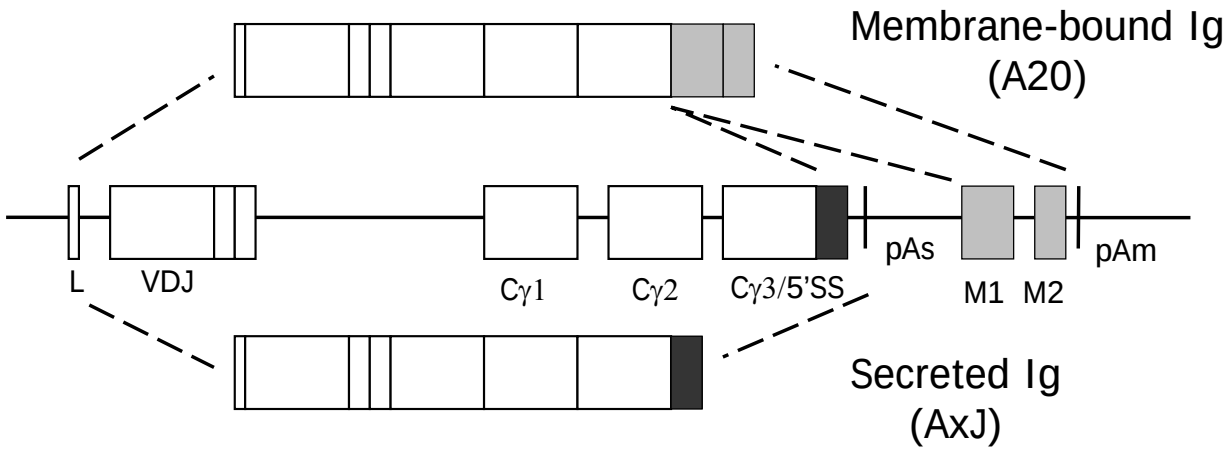


Figure 15. Alternative polyadenylation of the IgG2a gene.

The center of the figure depicts the genomic arrangement of the IgG2a gene after VDJ recombination has occurred where boxes represent protein coding exons and the line that connects the boxes represent introns and untranslated regions. When the 5' internal splice site (5'SS) is selected before cleavage/polyadenylation can occur at the secretory poly(A) site (pAs), the membrane form of Ig is created as a result of incorporation of the downstream membrane spanning exons (light gray) and use of the membrane bound associated poly(A) site (pAm). This occurs primarily in naïve and memory B-cells (e.g. A20 B-cell lymphoma). When the pAs is chosen for cleavage and polyadenylation, the complete Cγ3 exon is incorporated into the mature transcript (black box), thus excluding the incorporation of membrane spanning exons and producing the secretory form of Ig. This occurs primarily in plasma cells (e.g. AxJ hybridoma).

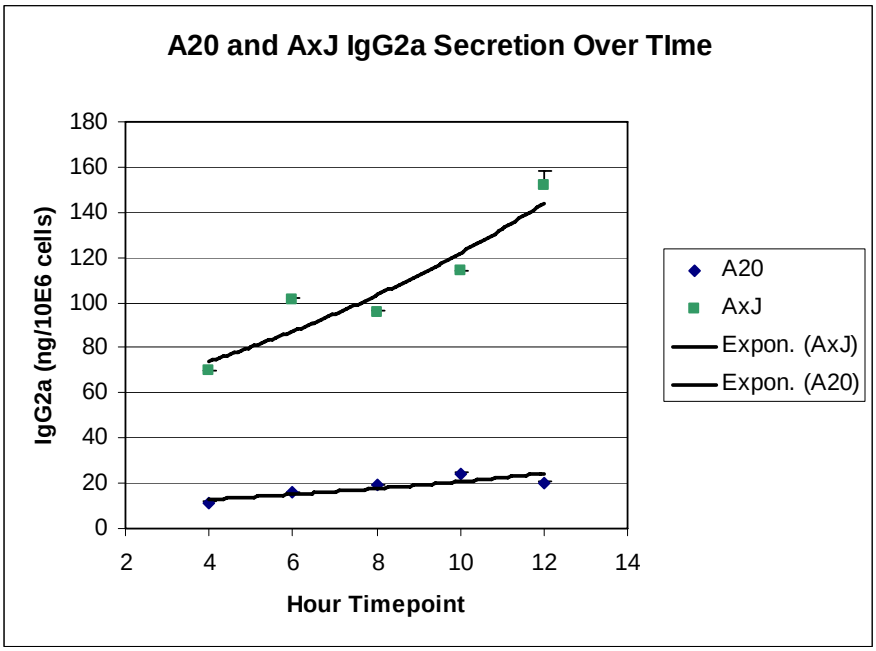
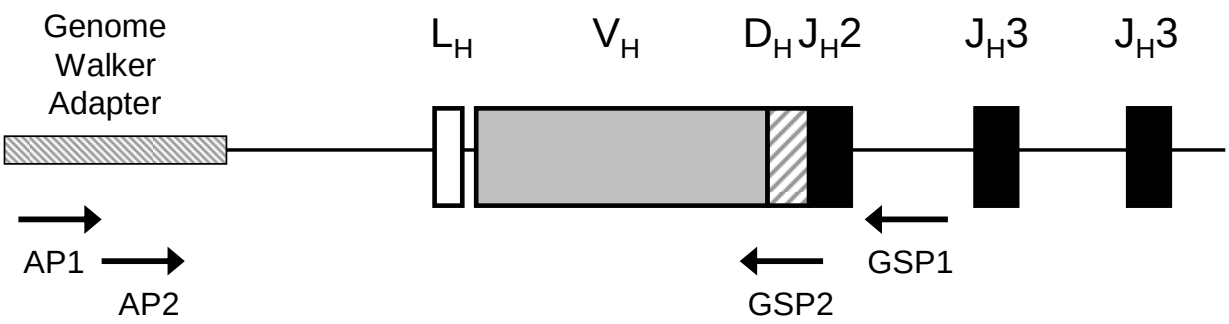


Figure 16. AxJ plasma cell hybridoma secretes Ig displaying a plasma cell phenotype.

ELISA was performed to measure IgG2a secretion by the A20 memory B-cell and the AxJ plasma cell hybridoma over time. Amounts of Ig per sample were measured against a standard curve and calculated per one million cells. An exponential curve was fit to the data. The measurement was taken in triplicate and SEM was calculated for each time point.

A



B

```
1   ACTATAGGGC ACGCGTGGTC GACGGCCCGG GCTGGTCTGT GTCACAGTGG
51  GGCCACTGTC TCAAGCTGCA AATCTTTTTA GTGCACAGGC TCTAATGTTA
101 CATCCATAGC CTCAACACAA GGTTCAGGGA TGAAGTATGG GATGAATTTC
151 CACAGACAAG ATGAGGACTT GGGCTTCAGA ATCCTGATTC CTGACCCAGA
201 TGTCCCTTCT TCTCCAGCAG GAGTAGGTGC TTATCTAATA TGTATCCTGC
251 TCATGAATAT GCAAATCCTG TGGGTCTACA GTGGTAAATA TAGGGTTGTC
301 TGCACGAAAC AGAAACCCTG AGATCGCTGC TCTCTTTACA GTTGCTGAAT
351 ACACAAGACC TCACCATGGC ATCGAACTTC ATCATGGTCT TCTTGACAAC
401 AACAGCTCCA GGTGAGGGGC TCACACTAAC AGCCTTGAGG TCTGGCCACA
451 CACATGGATG ACACTGACAC CAACCTTGCC TTTCTCTCCA CAGTTGTCCA
501 CTCCCCGGTC CAATTGCAAC AGTCCGGGCC TGACCTTGTG AAACCTGGGA
551 TGTCCGTGAA ATTGTCCTGT AAGACTTTGG GTTACAATTT CTCCGACAAG
601 TGGATTCACT GGATTAAACA GAAGCCTGGC CGAGGCCTTG AATGGGTTGG
651 AAGGATTGAT CCTTCTAACG GTGATACTGA CTATAATGTG GACTTCAAGA
701 CCCCGGCCAC ATTAACTGTT GACAGACCCT CCAACACAGC CTACTTAGAA
751 CTCAGTAACC TGACATCTGG GGACTCTGCG GTCTATTATT GTTCAATATC
801 GGGTGATTAT TCCGCCTGCG ACTATTGG
```

C

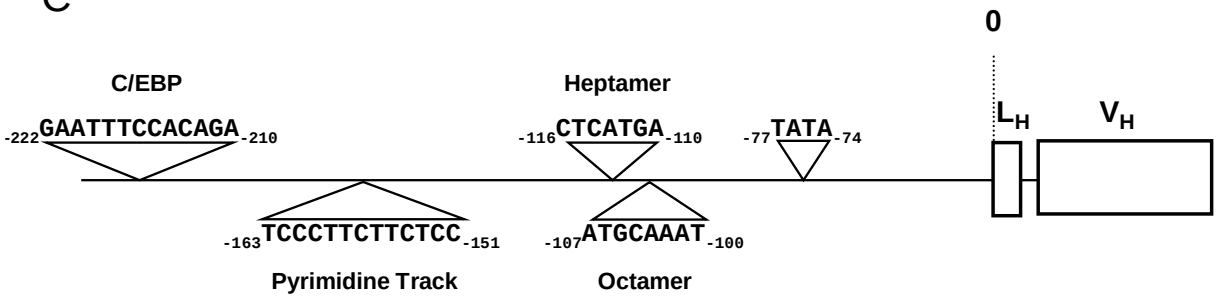


Figure 17. The A20 IgG2a promoter possesses elements common to most Ig promoters.

A. Illustrated is the 5'-RACE PCR strategy for obtaining the A20 IgG2a promoter sequence. The Genome Walker Adapter™ is ligated to the blunt-end of genomic DNA created by enzymatic digestion using PvuII. First round of PCR was performed using the linker-containing genomic library as template, the upstream-most adapter primer (AP1) and the downstream most gene-specific primer (GSP1). The second round of PCR was performed using the PCR product generated from the first round PCR as template, the nested adapter primer (AP2) and the nested gene-specific primer (GSP2). B. Shown is the entire PCR product generated from the nested second round of PCR. The underlined sequences are the heavy chain leader sequence and the heavy chain variable region, respectively. The sequence obtained from the 5'-RACE PCR for the leader and variable region are an exact match to the published A20 Ig sequence (39), assuring us that we had obtained the active Ig promoter in the A20 B-cell lymphoma line. C. The locations of the A20 IgG2a promoter elements are listed relative to the start codon contained in the heavy chain leader region.

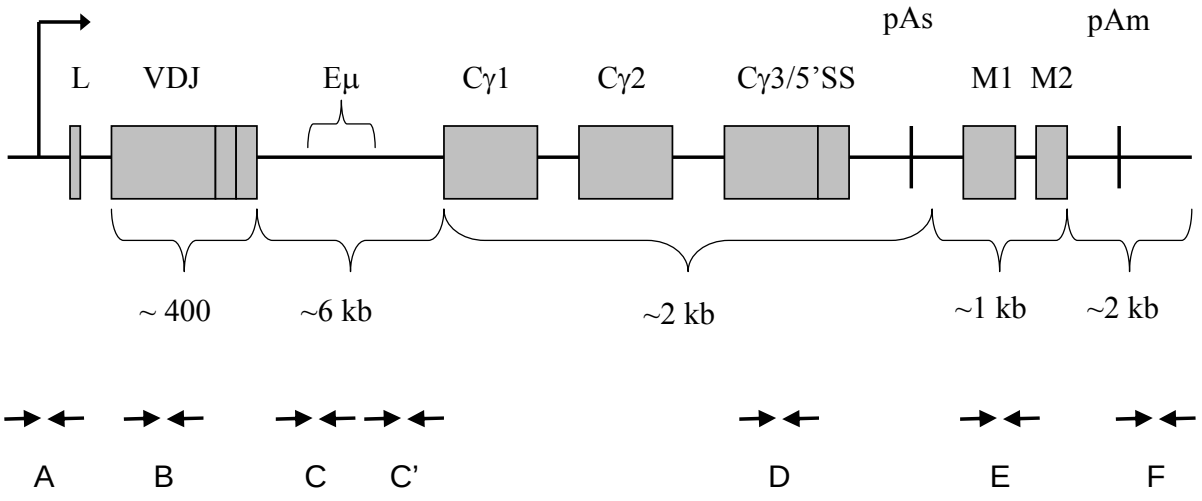
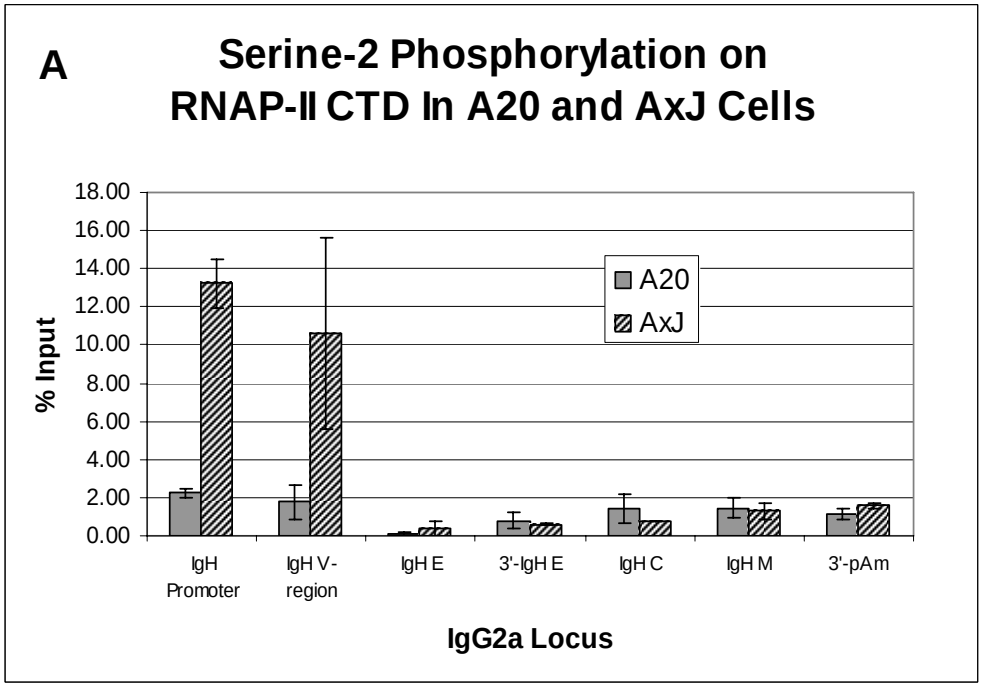


Figure 18. Location of primer pairs used in chromatin immunoprecipitation across the A20 IgG2a gene.

Illustrated is the exon arrangement of the A20 IgG2a genomic locus. Our approach to identify the phosphorylation state of RNAP-II CTD across the IgG2a genomic locus required the generation of multiple primer pairs. The location of many of the pairs was chosen based on important geographical features of the IgG2a genomic locus. Specifically, primer pair A targets the promoter of the A20 IgG2a gene, primer pair B targets the variable region, primer pairs C and C' target the area on and downstream of the intronic enhancer, Eμ. Locations of primer pairs D, E and F were chosen based on their locations relative to the secretory poly(A) site (pAs) and the membrane poly(A) site (pAm). Because we hypothesize that the phosphorylation state of the RNAP-II CTD and the association of mRNA processing factors may be quite dynamic at the 3'-end of the gene, we designed primers to flank both poly(A) sites. Therefore, primer pair D targets upstream of the pAs, primer pair E targets the membrane coding exon that lies between pAs and pAm, and primer pair F targets genomic sequence downstream of pAm. Indicated are approximate lengths of the elements and the vertical arrow identifies the general location of transcription initiation.


A
Serine-2 Phosphorylation on
RNAP-II CTD In A20 and AxJ Cells
% Input
18.00
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
A20
AxJ
IgH Promoter
IgH V-region
IgH E
3'-IgH E
IgH C
IgH M
3'-pAm
IgG2a Locus

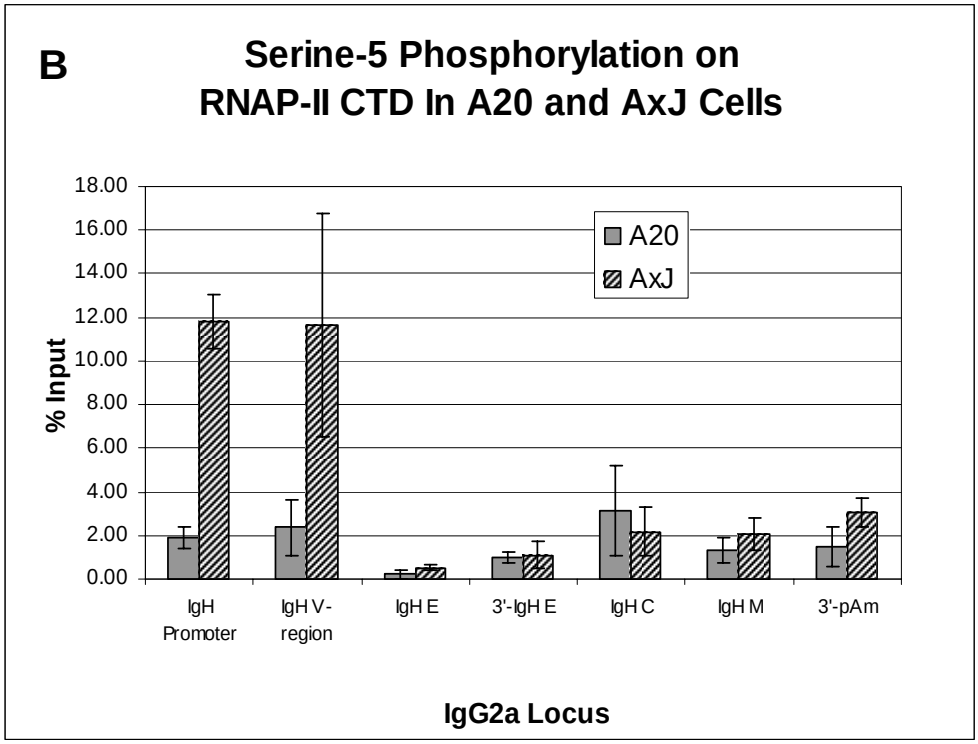

B
Serine-5 Phosphorylation on
RNAP-II CTD In A20 and AxJ Cells
% Input
18.00
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
A20
AxJ
IgH Promoter
IgH V-region
IgH E
3'-IgH E
IgH C
IgH M
3'-pAm
IgG2a Locus

Figure 19. Serine-2 and Serine-5 phosphorylation levels of RNAP-II CTD change across the IgG2a heavy chain genomic locus in B-lymphocytes

A. Serine-2 phosphorylation of RNAP-II CTD was measured by chromatin immunoprecipitation (ChIP). Shown are side-by-side comparisons of the level of Serine-2 phosphorylation of RNAP-II CTD between A20 memory B-cells and AxJ plasma cells at seven loci across the IgG2a heavy chain genomic locus. B. Serine-5 phosphorylation of RNAP-II CTD was measured by ChIP. Shown are side-by-side comparisons of the level of Serine-5 phosphorylation of RNAP-II CTD between A20 and AxJ cells at seven loci across the IgG2a heavy chain genomic locus. The amount of immunoprecipitated DNA was measured by QPCR and calculated as the percent input of total chromatin, as described in Materials and Methods. ChIP experiments at each locus were performed at least three times and the standard error of the mean (SEM) is shown.

APPENDIX A

Genes that increased expression in LPS-stimulated RAW macrophages[a]

Probe Set ID[b]	Gene Title	Gene Symbol	Ave. Signal Log Intensity[c]		Fold Increase[d]
			Untreated	LPS	
1427747_a_at	lipocalin 2	Lcn2	13.30	1846.23	138.81
1419427_at	colony stimulating factor 3 (granulocyte)	Csf3	30.17	4082.86	135.32
1416576_at	suppressor of cytokine signaling 3	Socs3	10.95	1313.06	119.93
1420380_at	chemokine (C-C motif) ligand 2	Ccl2	37.72	3350.24	88.82
1417263_at	prostaglandin-endoperoxide synthase 2	Ptgs2	49.32	4276.25	86.70
1418126_at	chemokine (C-C motif) ligand 5	Ccl5	59.95	5021.78	83.76
1449399_a_at	interleukin 1 beta	Il1b	48.33	4016.88	83.11
1427381_at	immunoresponsive gene 1	Irg1	60.50	4926.79	81.44
1421473_at	interleukin 1 alpha	Il1a	13.78	889.39	64.56
1449984_at	chemokine (C-X-C motif) ligand 2	Cxcl2	93.05	5570.33	59.86
1417314_at	histocompatibility 2, complement component factor B	H2-Bf	12.15	498.99	41.07
1450826_a_at	serum amyloid A 3	Saa3	56.95	2218.99	38.96
1436058_at	RIKEN cDNA 2510004L01 gene	2510004L01Rik	53.12	1920.50	36.15
1420393_at	nitric oxide synthase 2, inducible, macrophage	Nos2	10.01	348.13	34.79
1425951_a_at	C-type lectin, superfamily member 10	Clecsf10	10.57	355.51	33.62
1453196_a_at	2'-5' oligoadenylate synthetase-like 2	Oasl2	13.63	457.42	33.56
1418580_at	RIKEN cDNA 5830458K16 gene	5830458K16Rik	45.76	1522.32	33.27
1450783_at	interferon-induced protein with tetratricopeptide repeats 1	Ifit1	29.81	948.80	31.83
1417244_a_at	interferon regulatory factor 7	Irf7	95.21	2842.33	29.85
1419607_at	tumor necrosis factor	Tnf	146.81	4291.48	29.23
1420664_s_at	protein C receptor, endothelial	Procr	37.99	986.68	25.97
1419417_at	vascular endothelial growth factor C	Vegfc	23.22	595.02	25.63
1418240_at	guanylate nucleotide binding protein 2	Gbp2	10.01	246.68	24.65
1421207_at	leukemia inhibitory factor	Lif	11.35	273.97	24.13
1451924_a_at	endothelin 1	Edn1	10.07	233.15	23.15
1418835_at	pleckstrin homology-like domain, family A, member 1	Phlda1	24.61	568.55	23.10
1431591_s_at	interferon, alpha-inducible protein	G1p2	37.34	854.09	22.87
1419569_a_at	interferon-stimulated protein	Isg20	30.44	693.39	22.78
1419309_at	glycoprotein 38	Gp38	42.18	959.07	22.73
1417266_at	chemokine (C-C motif) ligand 6	Ccl6	11.48	256.64	22.36
1420549_at	guanylate nucleotide binding protein 1	Gbp1	14.10	305.49	21.66
1417483_at	expressed sequence AA408868	AA408868	47.59	1014.40	21.32
1423602_at	Tnf receptor-associated factor 1	Traf1	29.35	608.75	20.74
1419604_at	Z-DNA binding protein 1	Zbp1	10.01	205.55	20.54
1420697_at	solute carrier family 15, member 3	Slc15a3	122.99	2470.78	20.09

1424942_a_at	myelocytomatosis oncogene	Myc	22.43	441.76	19.69
1418293_at	interferon-induced protein with tetratricopeptide repeats 2	Ifit2	35.16	688.96	19.59
1448562_at	uridine phosphorylase 1	Upp1	10.01	192.69	19.26
1460251_at	tumor necrosis factor receptor superfamily, member 6	Tnfrsf6	11.17	213.44	19.11
1427348_at	cDNA sequence BC036563	BC036563	11.63	217.40	18.70
1419603_at	interferon, gamma-inducible protein 16	Ifi16	11.17	205.25	18.38
1424339_at	2'-5' oligoadenylate synthetase-like 1	Oasl1	117.50	2132.50	18.15
1451798_at	interleukin 1 receptor antagonist	Il1rn	312.22	5642.68	18.07
1449025_at	interferon-induced protein with tetratricopeptide repeats 3	Ifit3	20.50	363.82	17.75
1417494_a_at	ceruloplasmin	Cp	10.01	171.93	17.18
1421217_a_at	lectin, galactose binding, soluble 9	Lgals9	103.35	1745.08	16.88
1421551_s_at	interferon activated gene 202B	Ifi202b	168.18	2831.27	16.83
1420671_x_at	membrane-spanning 4-domains, subfamily A, member 4C	Ms4a4c	12.51	202.40	16.18
1425917_at	histocompatibility 28	H28	18.98	297.26	15.66
1435137_s_at	Mus musculus transcribed sequences		198.98	3112.49	15.64
1435331_at	expressed sequence AI447904	AI447904	13.79	209.29	15.18
1423954_at	complement component 3	C3	36.02	534.20	14.83
1418612_at	schlafen 1	Slfn1	10.01	147.90	14.78
1419149_at	serine (or cysteine) proteinase inhibitor, clade E, member 1	Serpine1	69.83	1026.58	14.70
1450297_at	interleukin 6	Il6	10.01	145.95	14.59
1421228_at	chemokine (C-C motif) ligand 7	Ccl7	11.01	157.03	14.26
1419714_at	programmed cell death 1 ligand 1	Pdcd1lg1	63.59	903.58	14.21
1434380_at	Mus musculus diabetic nephropathy-related gene 1 mRNA, partial sequence		16.56	231.31	13.96
1418392_a_at	guanylate nucleotide binding protein 3	Gbp3	10.45	144.73	13.85
1419598_at	membrane-spanning 4-domains, subfamily A, member 11	Ms4a11	121.83	1603.46	13.16
1422095_a_at	thymidylate kinase family LPS-inducible member	Tyki	14.46	187.36	12.96
1427005_at	serum-inducible kinase	Snk	149.57	1885.37	12.61
1448830_at	dual specificity phosphatase 1	Dusp1	80.86	1009.19	12.48
1418930_at	chemokine (C-X-C motif) ligand 10	Cxcl10	31.01	378.46	12.20
1422648_at	solute carrier family 7 (cationic amino acid transporter, y+ system), member 2	Slc7a2	10.48	126.53	12.07
1418191_at	ubiquitin specific protease 18	Usp18	85.31	988.66	11.59
1424921_at	RIKEN cDNA 2310015I10 gene	2310015I10Rik	140.94	1552.44	11.01
1418936_at	v-maf musculoaponeurotic fibrosarcoma oncogene family, protein F (avian)	Maff	10.01	106.50	10.64
1448951_at	tumor necrosis factor receptor superfamily, member 1b	Tnfrsf1b	71.53	752.06	10.51

1435792_at	component of Sp100-rs	Csprs	44.01	451.20	10.25
1443414_at	Mus musculus transcribed sequence with weak similarity to protein ref:NP_081764.1 (M.musculus) RIKEN cDNA 5730493B19 [Mus musculus]		53.81	545.32	10.13
1427736_a_at	chemokine (C-C motif) receptor-like 2	Ccrl2	39.24	397.26	10.12
1449591_at	caspase 4, apoptosis-related cysteine protease	Casp4	43.61	426.25	9.77
1418741_at	integrin beta 7	Itgb7	22.25	213.29	9.58
1450322_s_at	schlafen 3	Slfn3	100.16	948.94	9.47
1453472_a_at	SLAM family member 7	Slamf7	11.24	105.38	9.37
1434881_s_at	potassium channel tetramerisation domain containing 12	Kctd12	38.70	348.75	9.01
1423240_at	Rous sarcoma oncogene	Src	35.28	310.16	8.79
1451777_at	cDNA sequence BC013672	BC013672	10.75	94.04	8.75
1426276_at	RIKEN cDNA 9130009C22 gene	9130009C22Rik	79.53	683.51	8.59
1417925_at	chemokine (C-C motif) ligand 22	Ccl22	10.01	85.69	8.56
1450165_at	schlafen 2	Slfn2	213.36	1826.53	8.56
1436172_at	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4921508F21 product:similar to HISTOCOMPATIBILITY 2, CLASS II ANTIGEN E BETA [Mus musculus], full insert sequence		25.31	210.98	8.34
1421031_a_at	RIKEN cDNA 2310016C08 gene	2310016C08Rik	130.31	1072.77	8.23
1423555_a_at	RIKEN cDNA A430056A10 gene	A430056A10Rik	104.99	856.31	8.16
1421578_at	chemokine (C-C motif) ligand 4	Ccl4	558.94	4551.92	8.14
1425156_at	RIKEN cDNA 9830147J24 gene	9830147J24Rik	17.58	141.96	8.08
1419082_at	serine (or cysteine) proteinase inhibitor, clade B, member 2	Serpinb2	10.01	77.58	7.75
1420413_at	solute carrier family 7 (cationic amino acid transporter, y+ system), member 11	Slc7a11	78.87	610.36	7.74
1415922_s_at	MARCKS-like protein	Mlp	320.90	2464.17	7.68
1452521_a_at	urokinase plasminogen activator receptor	Plaur	147.07	1127.49	7.67
1427994_at	RIKEN cDNA F730004D16 gene	F730004D16Rik	66.95	489.45	7.31
1422177_at	interleukin 13 receptor, alpha 2	Il13ra2	26.51	193.66	7.31
1419561_at	chemokine (C-C motif) ligand 3	Ccl3	1077.22	7830.32	7.27
1449556_at	histocompatibility 2, T region locus 23	H2-T23	273.48	1984.72	7.26
1417516_at	DNA-damage inducible transcript 3	Ddit3	73.24	528.57	7.22
1451567_a_at	interferon activated gene 203	Ifi203	138.45	996.19	7.20
1448306_at	nuclear factor of kappa light chain gene enhancer in B-cells inhibitor, alpha	Nfkbia	564.80	4050.51	7.17
1416298_at	matrix metalloproteinase 9	Mmp9	44.38	315.55	7.11
1418847_at	arginase type II	Arg2	22.66	161.04	7.11

1438676_at	Mus musculus transcribed sequence with weak similarity to protein sp:P32456 (H.sapiens) GBP2_HUMAN Interferon-induced guanylate-binding protein 2 (Guanine nucleotide-binding protein 2)		10.01	70.66	7.06
1416593_at	glutaredoxin 1 (thioltransferase)	Glrx1	238.04	1668.01	7.01
1428027_at	Mus musculus BIC noncoding mRNA, complete sequence.		10.01	66.49	6.64
1422160_at	histocompatibility 2, T region locus 24	H2-T24	21.82	143.79	6.59
1416200_at	RIKEN cDNA 9230117N10 gene	9230117N10Rik	10.01	64.99	6.49
1455265_a_at	regulator of G-protein signaling 16	Rgs16	23.59	151.42	6.42
1449124_at	ral guanine nucleotide dissociation stimulator,-like 1	Rgl1	132.70	849.50	6.40
1448383_at	matrix metalloproteinase 14 (membrane-inserted)	Mmp14	21.12	134.65	6.38
1427102_at	schlafen 4	Slfn4	566.07	3579.11	6.32
1460415_a_at	tumor necrosis factor receptor superfamily, member 5	Tnfrsf5	36.70	231.06	6.30
1435226_at	RIKEN cDNA 4930534K13 gene	4930534K13Rik	109.18	683.37	6.26
1417961_a_at	tripartite motif protein 30	Trim30	111.87	691.61	6.18
1449450_at	prostaglandin E synthase	Ptges	57.46	353.18	6.15
1450214_at	adenosine A2b receptor	Adora2b	68.30	417.72	6.12
1418350_at	diphtheria toxin receptor	Dtr	25.68	156.82	6.11
1421304_at	killer cell lectin-like receptor, subfamily A, member 2	Klra2	20.45	123.84	6.06
1451426_at	DNA segment, Chr 11, Lothar Hennighausen 2, expressed	D11Lgp2e	69.85	419.74	6.01
1418186_at	glutathione S-transferase, theta 1	Gstt1	15.46	89.67	5.80
1419665_a_at	nuclear protein 1	Nupr1	102.46	592.07	5.78
1427592_at	Mus musculus similar to protocadherin 7 isoform c precursor;		56.44	324.86	5.76
1431705_a_at	mucolipin 2	Mcoln2	100.22	558.14	5.57
1420330_at	C-type (calcium dependent, carbohydrate recognition domain) lectin, superfamily member 9	Clecsf9	938.03	5149.77	5.49
1449363_at	activating transcription factor 3	Atf3	461.34	2498.09	5.41
1426243_at	cystathionase (cystathionine gamma-lyase)	Cth	36.91	199.66	5.41
1450672_a_at	three prime repair exonuclease 1	Trex1	115.25	610.99	5.30
1419609_at	chemokine (C-C motif) receptor 1	Ccr1	34.51	181.11	5.25
1424775_at	2'-5' oligoadenylate synthetase 1G	Oas1g	321.83	1688.37	5.25
1418626_a_at	clusterin	Clu	103.36	541.70	5.24
1419666_x_at	nuclear protein 1	Nupr1	94.77	489.04	5.16
1417023_a_at	fatty acid binding protein 4, adipocyte	Fabp4	38.51	196.96	5.11
1449731_s_at	nuclear factor of kappa light chain gene enhancer in B-cells inhibitor, alpha	Nfkbia	680.28	3472.31	5.10

1453721_a_at	solute carrier family 31, member 2	Slc31a2	162.72	827.85	5.09
1448377_at	secretory leukocyte protease inhibitor	Slpi	99.70	505.70	5.07
1418901_at	CCAAT/enhancer binding protein (C/EBP), beta	Cebpb	635.01	3218.70	5.07
1417876_at	Fc receptor, IgG, high affinity I	Fcgr1	105.80	535.68	5.06
1423006_at	proviral integration site 1	Pim1	47.72	240.72	5.04
1456494_a_at	tripartite motif protein 30-like	LOC209387	42.45	213.56	5.03
1417517_at	pleiomorphic adenoma gene-like 2	Plagl2	81.86	410.05	5.01
1439429_x_at	deltex 2 homolog (Drosophila)	Dtx2	39.27	196.68	5.01
1450957_a_at	sequestosome 1	Sqstm1	1099.48	5463.35	4.97
1450234_at	membrane-spanning 4-domains, subfamily A, member 6C	Ms4a6c	64.53	320.46	4.97
1448949_at	carbonic anhydrase 4	Car4	10.01	49.13	4.91
1417814_at	phospholipase A2, group V	Pla2g5	42.71	207.97	4.87
1418133_at	B-cell leukemia/lymphoma 3	Bcl3	46.74	226.19	4.84
1421326_at	colony stimulating factor 2 receptor, beta 1, low-affinity (granulocyte-macrophage)	Csf2rb1	75.58	364.73	4.83
1451655_at	schlafen 8	Slfn8	29.69	143.23	4.82
1416432_at	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3	Pfkfb3	143.63	690.68	4.81
1417813_at	inhibitor of kappaB kinase epsilon	Ikbke	40.32	192.08	4.76
1451544_at	similar to hypothetical protein FLJ10143	LOC213233	24.11	114.68	4.76
1438097_at	RAB20, member RAS oncogene family	Rab20	25.02	118.19	4.72
1450639_at	solute carrier family 28 (sodium-coupled nucleoside transporter), member 2	Slc28a2	32.51	153.37	4.72
1420499_at	GTP cyclohydrolase 1	Gch	209.91	984.80	4.69
1452196_a_at	NCK-associated protein 1	Nckap1	40.80	191.31	4.69
1421041_s_at	glutathione S-transferase, alpha 2 (Yc2)	Gsta2	10.01	46.85	4.68
1448239_at	heme oxygenase (decycling) 1	Hmox1	478.68	2235.75	4.67
1451821_a_at	nuclear antigen Sp100	Sp100	62.60	291.29	4.65
1416023_at	fatty acid binding protein 3, muscle and heart	Fabp3	32.41	150.67	4.65
1426971_at	Mus musculus NOD-derived CD11c +ve dendritic cells cDNA, clone:F630107D10 product:similar to UBIQUITIN-ACTIVATING ENZYME E1 [Homo sapiens]		38.84	180.32	4.64
1417371_at	pellino 1	Peli1	115.00	533.81	4.64
1450971_at	growth arrest and DNA-damage-inducible 45 beta	Gadd45b	222.81	1031.38	4.63
1424290_at	cDNA sequence BC010311	BC010311	161.37	736.61	4.56
1415972_at	myristoylated alanine rich protein kinase C substrate	Marcks	10.01	45.46	4.54
1417487_at	fos-like antigen 1	Fosl1	36.03	162.90	4.52
1424996_at	Mus musculus cDNA clone MGC:28609 IMAGE:4218551 complete cds		146.60	659.56	4.50

1418489_a_at	calcitonin receptor-like	Calcrl	29.26	131.51	4.49
1417193_at	superoxide dismutase 2, mitochondrial	Sod2	741.14	3288.67	4.44
1423268_at	integrin alpha 5 (fibronectin receptor alpha)	Itga5	80.48	356.91	4.44
1452087_at	RIKEN cDNA 2310046K10 gene	2310046K10Rik	124.20	548.33	4.41
1418826_at	membrane-spanning 4-domains, subfamily A, member 6B	Ms4a6b	60.78	266.09	4.38
1416975_at	signal transducing adaptor molecule (SH3 domain and ITAM motif) 2	Stam2	41.44	181.20	4.37
1451458_at	transmembrane protein 2	Tmem2	86.86	379.69	4.37
1418937_at	deiodinase, iodothyronine, type II	Dio2	11.19	48.52	4.33
1449453_at	bone marrow stromal cell antigen 1	Bst1	232.48	992.21	4.27
1427357_at	cytidine deaminase	Cda	46.17	196.44	4.25
1427313_at	prostaglandin I receptor (IP)	Ptgir	62.73	264.73	4.22
1419004_s_at	B-cell leukemia/lymphoma 2 related protein A1a	Bcl2a1a	663.02	2792.30	4.21
1418261_at	spleen tyrosine kinase	Syk	140.99	592.18	4.20
1422431_at	melanoma antigen, family E, 1	Magee1	36.26	152.24	4.20
1450027_at	syndecan 3	Sdc3	243.97	1020.84	4.18
1427689_a_at	TNFAIP3 interacting protein 1	Tnip1	315.20	1318.67	4.18
1429650_at	RIKEN cDNA 2310004N11 gene	2310004N11Rik	82.29	343.21	4.17
1416012_at	EH-domain containing 1	Ehd1	61.43	253.86	4.13
1448061_at	macrophage scavenger receptor 1	Msr1	334.23	1376.71	4.12
1416016_at	transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)	Tap1	104.05	428.25	4.12
1419721_at	interferon-gamma inducible gene, Puma-g	Pumag	104.54	428.20	4.10
1416239_at	argininosuccinate synthetase 1	Ass1	48.44	197.93	4.09
1451386_at	biliverdin reductase B (flavin reductase (NADPH))	Blvrb	222.13	893.80	4.02
1416381_a_at	peroxiredoxin 5	Prdx5	529.89	2113.04	3.99
1450884_at	CD36 antigen	Cd36	15.19	60.31	3.97
1418366_at	histone 2, H2aa1	Hist2h2aa1	94.97	377.03	3.97
1455581_x_at	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4921508F21 product:similar to HISTOCOMPATIBILITY 2, CLASS II ANTIGEN E BETA [Mus musculus]		64.79	256.59	3.96
1424831_at	copine II	Cpne2	24.62	97.31	3.95
1416521_at	selenoprotein W, muscle 1	Sepw1	86.41	341.33	3.95
1437111_at	Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length enriched library, clone:C230027N18 product:unknown EST		53.71	209.14	3.89

1434484_at	RIKEN cDNA 1100001G20 gene	1100001G20Rik	11.65	45.09	3.87
1418003_at	RIKEN cDNA 1190002H23 gene	1190002H23Rik	37.66	145.34	3.86
1431609_a_at	acid phosphatase 5, tartrate resistant	Acp5	79.95	306.12	3.83
1452536_s_at	immunoglobulin kappa chain variable 1 (V1)	Igk-V1	31.05	118.13	3.80
1449875_s_at	histocompatibility 2, T region locus 10	H2-T10	127.10	479.87	3.78
1450033_a_at	signal transducer and activator of transcription 1	Stat1	137.79	518.03	3.76
1453287_at	RIKEN cDNA 0610012A05 gene	0610012A05Rik	63.39	238.18	3.76
1416654_at	solute carrier family 31, member 2	Slc31a2	215.05	803.52	3.74
1416067_at	interferon-related developmental regulator 1	Ifrd1	419.87	1566.59	3.73
1424329_a_at	proline-rich Gla (G-carboxyglutamic acid) polypeptide 2	Prrg2	18.24	68.00	3.73
1416695_at	benzodiazepine receptor, peripheral	Bzrp	464.26	1729.85	3.73
1451161_a_at	EGF-like module containing, mucin-like, hormone receptor-like sequence 1	Emr1	504.91	1880.66	3.72
1419069_at	RAB guanine nucleotide exchange factor (GEF) 1	Rabgef1	21.19	78.71	3.71
1435174_at	RIKEN cDNA C230004D03 gene	C230004D03Rik	163.92	606.43	3.70
1448748_at	pleckstrin	Plek	940.25	3447.25	3.67
1418946_at	sialyltransferase 4A (beta-galactoside alpha-2,3-sialytransferase)	Siat4a	38.60	141.43	3.66
1419703_at	procollagen, type V, alpha 3	Col5a3	41.93	153.48	3.66
1456543_at	G protein-coupled receptor 73	Gpr73	24.84	90.47	3.64
1417542_at	ribosomal protein S6 kinase, polypeptide 2	Rps6ka2	130.68	474.51	3.63
1418401_a_at	dual specificity phosphatase 16	Dusp16	132.50	477.78	3.61
1423233_at	CCAAT/enhancer binding protein (C/EBP), delta	Cebpd	130.65	470.56	3.60
1426622_a_at	glutaminyl-peptide cyclotransferase (glutaminyl cyclase)	Qpct	14.30	51.34	3.59
1425078_x_at	RIKEN cDNA 4930565N07 gene	4930565N07Rik	229.44	817.42	3.56
1451564_at	cDNA sequence BC021340	BC021340	84.86	302.20	3.56
1450744_at	elongation factor RNA polymerase II 2	Ell2	365.25	1296.75	3.55
1421266_s_at	nuclear factor of kappa light chain gene enhancer in B-cells inhibitor, beta	Nfkbib	165.20	584.23	3.54
1448417_at	ninjurin 1	Ninj1	161.94	564.11	3.48
1455366_at	mitochondrial ribosomal protein L52	Mrpl52	11.03	38.35	3.48
1435176_a_at	inhibitor of DNA binding 2	Idb2	86.52	298.62	3.45
1426244_at	microtubule-associated protein, RP/EB family, member 2	Mapre2	109.21	376.04	3.44
1423605_a_at	transformed mouse 3T3 cell double minute 2	Mdm2	552.37	1890.49	3.42
1425902_a_at	nuclear factor of kappa light polypeptide gene enhancer in B-cells 2, p49/p100	Nfkb2	106.17	363.37	3.42
1425687_at	CASP8 and FADD-like apoptosis regulator	Cflar	30.44	103.97	3.42

1423686_a_at	RIKEN cDNA 1110020C13	1110020C13Rik	296.60	1013.01	3.42
1448575_at	interleukin 7 receptor	Il7r	247.54	845.22	3.41
1452078_a_at	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 2	Slc11a2	262.28	891.04	3.40
1416084_at	zinc finger protein 216	Zfp216	67.48	229.05	3.39
1422046_at	integrin alpha M	Itgam	164.83	556.47	3.38
1456309_x_at	LIM and SH3 protein 1	Lasp1	491.09	1657.21	3.37
1426906_at	interferon activated gene 205	Ifi205	311.02	1045.20	3.36
1418280_at	core promoter element binding protein	Copeb	106.37	356.28	3.35
1421308_at	carbonic anhydrase 13	Car13	35.69	119.39	3.35
1420361_at	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 1	Slc11a1	458.80	1530.87	3.34
1434547_at	carboxypeptidase D	Cpd	369.63	1225.01	3.31
1451338_at	nischarin	Nisch	35.97	118.89	3.31
1422704_at	glycerol kinase	Gyk	143.61	470.22	3.27
1420641_a_at	sulfide quinone reductase-like	Sqrdl	197.13	639.83	3.25
1421322_a_at	interferon dependent positive acting transcription factor 3 gamma	Isgf3g	83.82	272.03	3.25
1421065_at	Janus kinase 2	Jak2	88.26	286.39	3.24
1419647_a_at	immediate early response 3	Ier3	1174.58	3807.91	3.24
1416503_at	latexin	Lxn	107.46	346.20	3.22
1437161_x_at	RNA binding protein gene with multiple splicing	Rbpms	10.01	32.21	3.22
1448898_at	chemokine (C-C motif) ligand 9	Ccl9	1640.27	5271.37	3.21
1418641_at	lymphocyte cytosolic protein 2	Lcp2	455.61	1453.42	3.19
1434432_at	ligase III, DNA, ATP-dependent	Lig3	49.17	156.60	3.18
1451905_a_at	myxovirus (influenza virus) resistance 1	Mx1	10.39	33.02	3.18
1451828_a_at	fatty acid-Coenzyme A ligase, long chain 4	Facl4	169.02	535.85	3.17
1422851_at	high mobility group AT-hook 2	Hmga2	272.32	849.94	3.12
1424976_at	ras homolog gene family, member V	Arhv	13.38	41.73	3.12
1428942_at	metallothionein 2	Mt2	1376.28	4268.60	3.10
1460177_at	RIKEN cDNA 0610010E05	0610010E05Rik	430.18	1329.13	3.09
1426812_a_at	RIKEN cDNA 9130404D14	9130404D14Rik	573.89	1765.11	3.08
1416273_at	tumor necrosis factor, alpha-induced protein 2	Tnfaip2	317.66	966.63	3.04
1426696_at	low density lipoprotein receptor-related protein associated protein 1	Lrpap1	383.66	1161.41	3.03
1435518_at	RAS related protein 1b	Rap1b	784.04	2357.14	3.01
1452413_at	RIKEN cDNA C230081A13	C230081A13Rik	86.33	258.46	2.99
1415936_at	breast cancer anti-estrogen resistance 3	Bcar3	227.41	680.77	2.99
1421812_at	TAP binding protein	Tapbp	731.84	2165.64	2.96
1425974_a_at	tripartite motif protein 25	Trim25	576.34	1701.63	2.95
1431774_a_at	RIKEN cDNA 4930404J24	4930404J24Rik	18.53	54.48	2.94

1426278_at	RIKEN cDNA 2310061N23 gene	2310061N23Rik	356.04	1038.30	2.92
1423947_at	RIKEN cDNA 1110008P14 gene	1110008P14Rik	264.80	769.02	2.90
1423526_at	dead ringer homolog 2 (Drosophila)	Dri2	10.01	28.98	2.90
1460634_at	ral guanine nucleotide dissociation stimulator	Ralgds	118.74	342.96	2.89
1422924_at	tumor necrosis factor (ligand) superfamily, member 9	Tnfsf9	256.71	722.49	2.81
1423671_at	delta/notch-like EGF-related receptor	Dner	72.17	199.85	2.77
1423626_at	dystonin	Dst	66.31	183.54	2.77
1422650_a_at	RIO kinase 3 (yeast)	Riok3	318.44	880.93	2.77
1416901_at	Niemann Pick type C2	Npc2	2358.93	6469.15	2.74
1418842_at	hematopoietic cell specific Lyn substrate 1	Hcls1	432.67	1178.21	2.72
1418829_a_at	enolase 2, gamma neuronal	Eno2	305.70	827.21	2.71
1420888_at	Bcl2-like	Bcl2l	199.24	528.28	2.65
1425374_at	2'-5' oligoadenylate synthetase 3	Oas3	23.05	60.10	2.61
1422661_at	lectin, galactose binding, soluble 8	Lgals8	209.86	541.71	2.58
1417143_at	endothelial differentiation, lysophosphatidic acid G-protein-coupled receptor, 2	Edg2	28.53	73.58	2.58
1421098_at	expressed sequence AI586015	AI586015	822.89	2118.12	2.57
1449455_at	hemopoietic cell kinase	Hck	515.63	1325.77	2.57
1435660_at	RIKEN cDNA 5830484A20 gene	5830484A20Rik	102.99	261.42	2.54
1451941_a_at	Fc receptor, IgG, low affinity IIb	Fcgr2b	162.88	402.03	2.47
1420498_a_at	disabled homolog 2 (Drosophila)	Dab2	1032.34	2547.06	2.47
1420411_a_at	RIKEN cDNA 2610042N09 gene	2610042N09Rik	57.19	139.48	2.44
1419595_a_at	gamma-glutamyl hydrolase	Ggh	165.91	399.12	2.41
1418718_at	chemokine (C-X-C motif) ligand 16	Cxcl16	375.15	887.29	2.37
1425548_a_at	leukocyte specific transcript 1	Lst1	422.02	994.95	2.36
1452367_at	RIKEN cDNA 5830400N10 gene	5830400N10Rik	77.35	179.76	2.32

[a] Data are from three independent experiments and were analyzed as described in *Materials and Methods*. Genes shown were decreased an average of 2-fold or greater.
[b] Affymetrix probe ID number. Information on the probe set and the gene it targets can be accessed freely through http://www.affymetrix.com/analysis/index.affx.
[c] Average signal log intensity from three independent experiments.
[d] Fold decrease in average signal value observed in LPS-stimulated RAW macrophages compared to untreated, control.

APPENDIX B

Genes that decreased expression in LPS-stimulated RAW macrophages[a]

Probe Set ID[b]	Gene Title	Gene Symbol	Ave. Signal Log Intensity[c]		Fold Decrease[d]
			Untreated	LPS	
1448650_a_at	polymerase (DNA directed), epsilon	Pole	157.41	11.47	13.72
1419513_a_at	ect2 oncogene	Ect2	281.22	20.93	13.44
1452954_at	ubiquitin-conjugating enzyme E2C	Ube2c	1248.56	99.67	12.53
1450920_at	cyclin B2	Ccnb2	915.27	73.70	12.42
1450842_a_at	centromere autoantigen A	Cenpa	571.20	50.78	11.25
1426817_at	antigen identified by monoclonal antibody Ki 67	Mki67	1157.92	105.30	11.00
1423774_a_at	protein regulator of cytokinesis 1	Prc1	270.89	27.04	10.02
1430811_a_at	cell division cycle associated 1	Cdca1	356.16	37.03	9.62
1423463_a_at	DNA segment, Chr 2, ERATO Doi 750, expressed	D2Ertd750e	147.86	15.50	9.54
1449207_a_at	kinesin family member 20A	Kif20a	160.81	17.07	9.42
1448205_at	cyclin B1	Ccnb1	347.03	37.21	9.33
1427161_at	leucine, glutamic acid, lysine family 1 protein	Lek1	493.73	53.02	9.31
1439377_x_at	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 1	Elovl1	823.50	91.99	8.95
1433543_at	RIKEN cDNA 2900037I21 gene	2900037I21Rik	780.72	87.34	8.94
1416969_at	G two S phase expressed protein 1	Gtse1	108.71	12.52	8.68
1424766_at	cDNA sequence BC004701	BC004701	99.22	11.47	8.65
1424292_at	RIKEN cDNA 5830484J08 gene	5830484J08Rik	98.83	11.95	8.27
1423813_at	kinesin family member 22	Kif22	93.89	11.51	8.16
1416309_at	nucleolar and spindle associated protein 1	Nusap1	502.47	62.63	8.02
1416299_at	Shc SH2-domain binding protein 1	Shcbp1	401.32	51.72	7.76
1434695_at	RIKEN cDNA 2810047L02 gene	2810047L02Rik	168.22	22.03	7.63
1424278_a_at	baculoviral IAP repeat-containing 5	Birc5	445.25	58.43	7.62
1423847_at	RIKEN cDNA 2810406C15 gene	2810406C15Rik	289.75	39.40	7.35
1452242_at	RIKEN cDNA 1200008O12 gene	1200008O12Rik	613.24	83.59	7.34
1439040_at	cDNA sequence BC049989	BC049989	102.97	14.21	7.24
1422603_at	RIKEN cDNA C730049F20 gene	C730049F20Rik	156.62	21.64	7.24
1416258_at	thymidine kinase 1	Tk1	505.91	71.04	7.12
1429171_a_at	Mus musculus adult male hypothalamus cDNA, RIKEN full-length enriched library, clone:A230105F24 product:hypothetical protein, full insert sequence		180.95	25.64	7.06
1438009_at	histone 1, H2ae	Hist1h2ae	2497.44	354.76	7.04

1416664_at	cell division cycle 20 homolog (S. cerevisiae)	Cdc20	732.14	106.18	6.90
1422460_at	MAD2 (mitotic arrest deficient, homolog)-like 1 (yeast)	Mad2l1	549.55	80.24	6.85
1417822_at	DNA segment, Chr 17, human D6S56E 5	D17H6S56E-5	1268.11	187.35	6.77
1422462_at	RIKEN cDNA 2700084L22 gene	2700084L22Rik	228.33	33.99	6.72
1434850_at	hypothetical protein D030034H08	D030034H08	425.41	63.66	6.68
1448226_at	ribonucleotide reductase M2	Rrm2	499.80	74.87	6.68
1422252_a_at	cell division cycle 25 homolog C (S. cerevisiae)	Cdc25c	76.32	11.47	6.65
1437405_a_at	insulin-like growth factor binding protein 4	Igfbp4	884.06	134.54	6.57
1438852_x_at	minichromosome maintenance deficient 6 (MIS5 homolog, S. pombe) (S. cerevisiae)	Mcm6	335.31	51.16	6.55
1416030_a_at	Mus musculus adult male thymus cDNA, RIKEN full-length enriched library, clone:5830410A10 product:mini chromosome maintenance deficient 7 (S. cerevisiae), full insert sequence		791.17	121.97	6.49
1416774_at	wee 1 homolog (S. pombe)	Wee1	107.78	16.66	6.47
1416558_at	maternal embryonic leucine zipper kinase	Melk	220.26	34.25	6.43
1454694_a_at	topoisomerase (DNA) II alpha	Top2a	1124.03	175.71	6.40
1417450_a_at	transforming, acidic coiled-coil containing protein 3	Tacc3	135.08	21.23	6.36
1450886_at	germ cell-specific gene 2	Gsg2	85.02	13.42	6.34
1448191_at	polo-like kinase (Drosophila)	Plk	401.58	63.93	6.28
1452458_s_at	RIKEN cDNA 2410005L11 gene	2410005L11Rik	168.81	26.94	6.27
1433892_at	sperm associated antigen 5	Spag5	84.23	13.45	6.26
1452314_at	kinesin family member 11	Kif11	185.97	29.70	6.26
1417910_at	cyclin A2	Ccna2	604.89	96.76	6.25
1436186_at	RIKEN cDNA 4432406C08 gene	4432406C08Rik	144.78	23.57	6.14
1419620_at	pituitary tumor-transforming 1	Pttg1	102.77	16.91	6.08
1452040_a_at	cell division cycle associated 3	Cdca3	375.67	62.11	6.05
1435306_a_at	Mus musculus transcribed sequences		216.37	36.87	5.87
1448314_at	cell division cycle 2 homolog A (S. pombe)	Cdc2a	943.05	161.31	5.85
1449171_at	Ttk protein kinase	Ttk	202.44	35.70	5.67
1448627_s_at	DNA segment, Chr 14, ERATO Doi 732, expressed	D14Ertd732e	691.19	122.80	5.63
1450496_a_at	RIKEN cDNA 2810433K01 gene	2810433K01Rik	75.21	13.42	5.60
1452073_at	RIKEN cDNA 6720460F02 gene	6720460F02Rik	207.58	37.16	5.59
1450157_a_at	hyaluronan mediated motility receptor (RHAMM)	Hmmr	178.72	32.85	5.44
1451970_at	RIKEN cDNA E330036I19 gene	E330036I19Rik	145.90	26.95	5.41
1451246_s_at	aurora kinase B	Aurkb	167.28	31.22	5.36
1425753_a_at	uracil-DNA glycosylase	Ung	178.59	33.34	5.36
1449060_at	kinesin family member 2C	Kif2c	175.75	33.40	5.26
1428870_at	nucleolar and coiled-body phosphoprotein 1	Nolc1	60.30	11.47	5.26

1416757_at	RIKEN cDNA 2310031L18 gene	2310031L18Rik	212.86	41.15	5.17
1417926_at	RIKEN cDNA 5830426I05 gene	5830426I05Rik	212.96	41.41	5.14
1424223_at	RIKEN cDNA 1700020C11 gene	1700020C11Rik	75.09	14.66	5.12
1415945_at	minichromosome maintenance deficient 5, cell division cycle 46 (S. cerevisiae)	Mcm5	554.13	110.59	5.01
1451019_at	cathepsin F	Ctsf	132.79	26.83	4.95
1454011_a_at	replication protein A2	Rpa2	195.34	39.67	4.92
1418481_at	expressed sequence AW209059	AW209059	249.52	51.00	4.89
1449877_s_at	kinesin family member C5A	Kifc5a	165.84	33.92	4.89
1428669_at	brain expressed myelocytomatosis oncogene	Bmyc	246.52	50.47	4.88
1460713_at	Mus musculus mRNA similar to putative c-Myc-responsive (cDNA clone MGC:54855 IMAGE:5388297), complete cds		104.24	21.60	4.83
1449460_at	ankyrin repeat and SOCS box-containing protein 13	Asb13	64.59	13.50	4.78
1428105_at	RIKEN cDNA 2610005B21 gene	2610005B21Rik	65.53	13.74	4.77
1452241_at	RIKEN cDNA 2810429C13 gene	2810429C13Rik	489.93	103.19	4.75
1417327_at	caveolin 2	Cav2	86.99	18.50	4.70
1438320_s_at	Mus musculus adult male thymus cDNA, RIKEN full-length enriched library, clone:5830410A10 product:mini chromosome maintenance deficient 7 (S. cerevisiae), full insert sequence		1330.52	283.94	4.69
1426631_at	RIKEN cDNA C330017I15 gene	C330017I15Rik	225.09	48.15	4.67
1416961_at	budding uninhibited by benzimidazoles 1 homolog, beta (S. cerevisiae)	Bub1b	582.57	124.69	4.67
1424692_at	RIKEN cDNA 2810055F11 gene	2810055F11Rik	89.61	19.20	4.67
1452917_at	replication factor C (activator 1) 5	Rfc5	408.26	87.63	4.66
1416746_at	H2A histone family, member X	H2afx	841.98	183.42	4.59
1418264_at	SoxLZ/Sox6 leucine zipper binding protein in testis	Solt	163.66	35.72	4.58
1420477_at	nucleosome assembly protein 1-like 1	Nap1l1	151.18	33.15	4.56
1426473_at	RIKEN cDNA 5330419I01 gene	5330419I01Rik	304.01	68.17	4.46
1417971_at	RIKEN cDNA 2310014H01 gene	2310014H01Rik	337.32	75.88	4.45
1453107_s_at	phosphatidylethanolamine binding protein	Pbp	265.37	59.86	4.43
1449061_a_at	DNA primase, p49 subunit	Prim1	110.41	24.92	4.43
1428074_at	Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310037P21 product:BRAIN SPECIFIC BINDING PROTEIN homolog [Rattus norvegicus], full insert sequence		107.88	24.56	4.39
1436824_x_at	ring finger protein 26	Rnf26	450.39	102.63	4.39
1424895_at	RIKEN cDNA 6230410J09	6230410J09Rik	266.26	61.57	4.32

	gene				
1427105_at	RIKEN cDNA 2610510J17 gene	2610510J17Rik	434.50	100.69	4.32
1456393_at	programmed cell death 4	Pdcd4	215.41	49.99	4.31
1422430_at	fidgetin-like 1	Fignl1	244.93	57.02	4.30
1416889_at	troponin I, skeletal, fast 2	Tnni2	54.96	12.80	4.29
1448134_at	paladin	Pald	65.49	15.40	4.25
1422814_at	calmodulin binding protein 1	Calmbp1	230.94	54.43	4.24
1417019_a_at	cell division cycle 6 homolog (S. cerevisiae)	Cdc6	99.15	23.37	4.24
1417299_at	NIMA (never in mitosis gene a)-related expressed kinase 2	Nek2	129.47	30.52	4.24
1438390_s_at	pituitary tumor-transforming 1	Pttg1	2428.90	573.25	4.24
1452534_a_at	high mobility group box 2	Hmgb2	602.13	142.49	4.23
1435194_at	heat shock protein 4	Hspa4	81.61	19.32	4.22
1450862_at	RAD54 like (S. cerevisiae)	Rad54l	280.91	66.65	4.21
1452606_at	RIKEN cDNA 2610034E18 gene	2610034E18Rik	48.24	11.47	4.20
1424107_at	kinesin family member 18A	Kif18a	87.87	20.96	4.19
1424629_at	breast cancer 1	Brca1	263.22	63.07	4.17
1417647_at	sorting nexin 5	Snx5	339.33	81.32	4.17
1448635_at	SMC2 structural maintenance of chromosomes 2-like 1 (yeast)	Smc2l1	650.21	156.82	4.15
1429295_s_at	thyroid hormone receptor interactor 13	Trip13	295.08	71.41	4.13
1455990_at	kinesin family member 23	Kif23	408.70	100.00	4.09
1448113_at	stathmin 1	Stmn1	50.42	12.36	4.08
1448441_at	CDC28 protein kinase 1	Cks1	1016.21	250.72	4.05
1451849_a_at	lamin B2	Lmnb2	46.17	11.47	4.02
1418281_at	RAD51 homolog (S. cerevisiae)	Rad51	302.50	75.23	4.02
1448777_at	minichromosome maintenance deficient 2 mitotin (S. cerevisiae)	Mcm2	545.91	136.68	3.99
1450155_at	integrin alpha 4	Itga4	195.73	49.16	3.98
1418585_at	cyclin H	Ccnh	47.72	12.02	3.97
1420028_s_at	minichromosome maintenance deficient 3 (S. cerevisiae)	Mcm3	717.01	182.52	3.93
1451358_a_at	Rac GTPase-activating protein 1	Racgap1	519.81	132.39	3.93
1439269_x_at	Mus musculus adult male thymus cDNA, RIKEN full-length enriched library, clone:5830410A10 product:mini chromosome maintenance deficient 7 (S. cerevisiae), full insert sequence		161.99	41.27	3.93
1419270_a_at	deoxyuridine triphosphatase	Dutp	238.06	61.09	3.90
1450019_at	chemokine (C-X3-C) receptor 1	Cx3cr1	231.80	59.97	3.87
1424321_at	replication factor C (activator 1) 4	Rfc4	221.11	57.26	3.86
1428483_a_at	RIKEN cDNA 2610039C10 gene	2610039C10Rik	122.73	31.85	3.85
1450070_s_at	p21 (CDKN1A)-activated kinase 1	Pak1	254.68	66.12	3.85
1426653_at	minichromosome maintenance deficient 3 (S. cerevisiae)	Mcm3	347.16	90.23	3.85
1419153_at	RIKEN cDNA 2810417H13 gene	2810417H13Rik	996.68	259.52	3.84

1435054_at	hypothetical protein 6820428D13	6820428D13	107.21	28.04	3.82
1436707_x_at	RIKEN cDNA A730011O11 gene	A730011O11Rik	586.56	154.78	3.79
1418969_at	S-phase kinase-associated protein 2 (p45)	Skp2	209.22	55.53	3.77
1455730_at	hepatoma up-regulated protein	Hurp	162.36	43.44	3.74
1424156_at	retinoblastoma-like 1 (p107)	Rbl1	319.02	86.58	3.68
1428029_a_at	RIKEN cDNA C530002L11 gene	C530002L11Rik	293.92	79.82	3.68
1451103_at	RIKEN cDNA 9430093H08 gene	9430093H08Rik	309.50	84.12	3.68
1435697_a_at	pleckstrin homology, Sec7 and coiled-coil domains, binding protein	Pscdbp	146.56	39.90	3.67
1419998_at	RIKEN cDNA C330027C09 gene	C330027C09Rik	41.97	11.47	3.66
1438386_x_at	gamma-glutamyl carboxylase	Ggcx	302.52	83.12	3.64
1439381_x_at	expressed sequence AI504298	AI504298	341.43	94.01	3.63
1415810_at	nuclear protein 95	Np95	187.34	51.83	3.61
1448710_at	chemokine (C-X-C motif) receptor 4	Cxcr4	56.99	15.81	3.60
1418561_at	splicing factor 3b, subunit 1	Sf3b1	59.66	16.64	3.59
1423795_at	splicing factor proline/glutamine rich (polypyrimidine tract binding protein associated)	Sfpq	202.72	56.90	3.56
1422016_a_at	centromere autoantigen H	Cenph	67.63	19.04	3.55
1415829_at	lamin B receptor	Lbr	923.80	260.93	3.54
1419397_at	polymerase (DNA directed), alpha 1	Pola1	204.72	58.38	3.51
1418167_at	transcription factor AP-4 (activating enhancer-binding protein 4)	Tfap4	82.55	23.86	3.46
1417586_at	timeless homolog (Drosophila)	Timeless	181.73	52.66	3.45
1422513_at	cyclin F	Ccnf	136.05	39.49	3.45
1416155_at	high mobility group box 3	Hmgb3	309.28	89.78	3.45
1437302_at	adrenergic receptor, beta 2	Adrb2	82.48	23.99	3.44
1424033_at	splicing factor, arginine/serine-rich 7	Sfrs7	167.32	48.71	3.43
1423524_at	RIKEN cDNA 2700091H24 gene	2700091H24Rik	89.43	26.29	3.40
1435737_a_at	nuclear distribution gene E homolog 1 (A nidulans)	Nde1	215.44	63.46	3.39
1424511_at	serine/threonine kinase 6	Stk6	412.30	121.48	3.39
1422503_s_at	ADP-ribosyltransferase (NAD+; poly (ADP-ribose) polymerase) 1	Adprt1	637.82	188.29	3.39
1424118_a_at	RIKEN cDNA 2600017H08 gene	2600017H08Rik	189.38	55.91	3.39
1454636_at	chromobox homolog 5 (Drosophila HP1a)	Cbx5	348.57	104.34	3.34
1452508_x_at	RIKEN cDNA 2610009E16 gene	2610009E16Rik	51.47	15.42	3.34
1418872_at	ATP-binding cassette, sub-family B (MDR/TAP), member 1B	Abcb1b	178.06	53.52	3.33
1452829_at	carbamoyl-phosphate synthetase 2, aspartate transcarbamylase, and dihydroorotase	Cad	398.96	120.15	3.32
1450876_at	complement component factor h	Cfh	38.07	11.47	3.32

1434856_at	hypothetical protein A130096K20	A130096K20	94.36	28.47	3.31
1420928_at	beta galactoside alpha 2,6 sialyltransferase 1	St6gal1	199.23	60.18	3.31
1437050_s_at	DNA segment, Chr 1, ERATO Doi 396, expressed	D1Ertd396e	73.18	22.16	3.30
1448127_at	ribonucleotide reductase M1	Rrm1	783.22	237.17	3.30
1454699_at	sestrin 1	Sesn1	155.51	47.14	3.30
1452199_at	RIKEN cDNA 2700094F01 gene	2700094F01Rik	83.59	25.37	3.30
1423714_at	ASF1 anti-silencing function 1 homolog B (S. cerevisiae)	Asf1b	477.88	145.68	3.28
1421933_at	chromobox homolog 5 (Drosophila HP1a)	Cbx5	37.59	11.51	3.27
1424173_at	RIKEN cDNA 2810475A17 gene	2810475A17Rik	71.47	21.90	3.26
1423045_at	nuclear cap binding protein subunit 2	Ncbp2	68.92	21.32	3.23
1424716_at	RIKEN cDNA 0610039N19 gene	0610039N19Rik	83.40	26.00	3.21
1435323_a_at	O-acyltransferase (membrane bound) domain containing 1	Oact1	158.53	49.48	3.20
1426917_s_at	RIKEN cDNA 4833415E20 gene	4833415E20Rik	52.34	16.35	3.20
1424991_s_at	thymidylate synthase	Tyms	210.49	65.96	3.19
1425798_a_at	RecQ protein-like	Recql	63.08	19.89	3.17
1426767_at	RIKEN cDNA 3230401M21 gene	3230401M21Rik	133.38	42.20	3.16
1449065_at	cytosolic acyl-CoA thioesterase 1	Cte1	112.39	36.13	3.13
1424852_at	myocyte enhancer factor 2C	Mef2c	543.54	175.68	3.09
1416802_a_at	RIKEN cDNA 2610036L13 gene	2610036L13Rik	327.94	107.43	3.05
1451348_at	expressed sequence R75183	R75183	46.36	15.19	3.05
1417939_at	RAD51 associated protein 1	Rad51ap1	34.88	11.47	3.04
1426258_at	sortilin-related receptor, LDLR class A repeats-containing	Sorl1	235.80	77.79	3.03
1424046_at	budding uninhibited by benzimidazoles 1 homolog (S. cerevisiae)	Bub1	754.46	249.16	3.03
1427496_at	expressed sequence AI851464	AI851464	37.05	12.26	3.02
1450692_at	kinesin family member 4	Kif4	170.36	56.71	3.00
1416152_a_at	splicing factor, arginine/serine-rich 3 (SRp20)	Sfrs3	88.25	29.99	2.94
1420142_s_at	proliferation-associated 2G4	Pa2g4	157.14	54.34	2.89
1452659_at	DEK oncogene (DNA binding)	Dek	1057.96	368.17	2.87
1418004_a_at	RIKEN cDNA 1810009M01 gene	1810009M01Rik	150.74	52.48	2.87
1424407_s_at	neuronal pentraxin receptor	Nptxr	200.84	70.46	2.85
1433512_at	Friend leukemia integration 1	Fli1	259.09	91.43	2.83
1423667_at	methionine adenosyltransferase II, alpha	Mat2a	536.00	189.16	2.83
1415772_at	nucleolin	Ncl	88.15	31.21	2.82
1415997_at	thioredoxin interacting protein	Txnip	489.90	174.84	2.80
1454971_x_at	Mus musculus transcribed sequences		161.90	57.80	2.80
1415860_at	karyopherin (importin) alpha 2	Kpna2	1227.06	441.28	2.78
1427233_at	serologically defined colon cancer antigen 33	Sdccag33	110.90	40.05	2.77
1436349_at	Mus musculus 11 days embryo whole body cDNA, RIKEN		349.78	126.42	2.77

	full-length enriched library, clone:2700094K13 product:unknown EST, full insert sequence				
1423074_at	lectin, mannose-binding 2	Lman2	96.47	35.48	2.72
1427469_at	RIKEN cDNA 9630002H22 gene	9630002H22Rik	37.18	13.76	2.70
1418027_at	exonuclease 1	Exo1	58.36	21.97	2.66
1423521_at	lamin B1	Lmnb1	328.71	125.34	2.62
1418288_at	lipin 1	Lpin1	200.58	77.02	2.60
1436884_x_at	Ewing sarcoma homolog	Ewsh	255.57	98.42	2.60
1425742_a_at	transforming growth factor beta 1 induced transcript 4	Tgfb1i4	171.15	65.93	2.60
1418656_at	LSM5 homolog, U6 small nuclear RNA associated (S. cerevisiae)	Lsm5	196.00	75.69	2.59
1418380_at	telomeric repeat binding factor 1	Terf1	75.20	29.20	2.58
1425922_a_at	neuroblastoma myc-related oncogene 1	Nmyc1	516.97	201.88	2.56
1452210_at	RIKEN cDNA E130315B21 gene	E130315B21Rik	242.46	94.93	2.55
1420707_a_at	TRAF-interacting protein	Traip	29.11	11.47	2.54
1450112_a_at	growth arrest specific 2	Gas2	36.53	14.62	2.50
1438169_a_at	RIKEN cDNA 6030440G05 gene	6030440G05Rik	444.69	178.91	2.49
1423440_at	RIKEN cDNA 1110001A07 gene	1110001A07Rik	185.75	76.71	2.42
1431506_s_at	Mus musculus transcribed sequence with strong similarity to protein ref:NP_006338.1 (H.sapiens) peptidyl prolyl isomerase H (cyclophilin H) [Homo sapiens]		209.52	87.13	2.40
1423620_at	RIKEN cDNA 2610528M18 gene	2610528M18Rik	176.30	73.86	2.39
1422438_at	epoxide hydrolase 1, microsomal	Ephx1	438.98	184.54	2.38
1448673_at	poliovirus receptor-related 3	Pvrl3	183.87	77.69	2.38
1428788_at	RIKEN cDNA 1700012G19 gene	1700012G19Rik	167.13	70.77	2.36
1423064_at	DNA methyltransferase 3A	Dnmt3a	267.42	117.46	2.28
1422997_s_at	mitochondrial acyl-CoA thioesterase 1	Mte1	134.07	58.96	2.27
1434239_at	expressed sequence AA408556	AA408556	31.57	13.93	2.27
1436448_a_at	prostaglandin-endoperoxide synthase 1	Ptgs1	24.90	11.47	2.17

BIBLIOGRAPHY

1. **Abdel Wahab, N., J. Gibbs, and R. M. Mason.** 1998. Regulation of gene expression by alternative polyadenylation and mRNA instability in hyperglycaemic mesangial cells. Biochem J **336:**405-11.

2. **Ahn, S. H., M. Kim, and S. Buratowski.** 2004. Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. Mol Cell **13:**67-76.

3. **Appleby, S. B., A. Ristimaki, K. Neilson, K. Narko, and T. Hla.** 1994. Structure of the human cyclo-oxygenase-2 gene. Biochem J **302:**723-7.

4. **Aravind, L.** 1999. An evolutionary classification of the metallo-beta-lactamase fold proteins. In Silico Biol **1:**69-91.

5. **Arhin, G. K., M. Boots, P. S. Bagga, C. Milcarek, and J. Wilusz.** 2002. Downstream sequence elements with different affinities for the hnRNP H/H' protein influence the processing efficiency of mammalian polyadenylation signals. Nucleic Acids Res **30:**1842-50.

6. **Bagga, P., G. Arhin, and J. Wilusz.** 1998. DSEF-1 is a member of the hnRNP H family of RNA-binding proteins and stimulates pre-mRNA cleavage and polyadenylation in vitro. Nucleic Acids Research **26:**5343-50.

7. **Bagga, P. S., L. P. Ford, F. Chen, and J. Wilusz.** 1995. The G-rich auxiliary downstream element has distinct sequence and position requirements and mediates efficient 3' end pre-mRNA processing through a trans-acting factor. Nucleic Acids Res **23:**1625-31.

8. **Ballard, D. W., and A. Bothwell.** 1986. Mutational analysis of the immunoglobulin heavy chain promoter region. Proc Natl Acad Sci U S A **83:**9626-30.

9. **Beaudoing, E., S. Freier, J. R. Wyatt, J. M. Claverie, and D. Gautheret.** 2000. Patterns of variant polyadenylation signal usage in human genes. Genome Res **10:**1001-10.

10. **Beaudoing, E., and D. Gautheret.** 2001. Identification of alternate polyadenylation sites and analysis of their tissue distribution using EST data. Genome Res **11:**1520-6.

11. **Bienroth, S., E. Wahle, C. Suter-Crazzolara, and W. Keller.** 1991. Purification of the cleavage and polyadenylation factor involved in the 3'-processing of messenger RNA precursors. Journal of Biological Chemistry **266:**19768-19776.

12. **Bird, G., D. A. Zorio, and D. L. Bentley.** 2004. RNA Polymerase II Carboxy-Terminal Domain Phosphorylation Is Required for Cotranscriptional Pre-mRNA Splicing and 3'-End Formation. Mol Cell Biol **24:**8963-9.

13. **Bodnar, K. A., N. V. Serbina, and J. L. Flynn.** 2001. Fate of Mycobacterium tuberculosis within murine dendritic cells. Infect Immun **69:**800-9.

14. **Bourrouillou, G., P. Colombies, P. Blanc, J. Grozdea, and F. Pontonnier.** 1978. [Chromosome aberrations in sterile males. Study of 241 cases]. Nouv Presse Med **7:**3777.

15. **Bruce, S. R., R. W. Dingle, and M. L. Peterson.** 2003. B-cell and plasma-cell splicing differences: a potential role in regulated immunoglobulin RNA processing. Rna **9:**1264-73.

16. **Burke, B., A. Giannoudis, K. P. Corke, D. Gill, M. Wells, L. Ziegler-Heitbrock, and C. E. Lewis.** 2003. Hypoxia-induced gene expression in human macrophages: implications for ischemic tissues and hypoxia-regulated gene therapy. Am J Pathol **163:**1233-43.

17. **Burtis, K. C., and B. S. Baker.** 1989. Drosophila doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. Cell **56:**997-1010.

18. **Calvo, O., and J. L. Manley.** 2001. Evolutionarily conserved interaction between CstF-64 and PC4 links transcription polyadenylation, and termination. Mol Cell **7:**1013-1023.

19. **Caputi, M., and A. M. Zahler.** 2001. Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) H/H'/F/2H9 family. J Biol Chem **276:**43850-9. Epub 2001 Sep 24.

20. **Chen, C. Y., T. M. Chen, and A. B. Shyu.** 1994. Interplay of two functionally and structurally distinct domains of the c-fos AU-rich element specifies its mRNA-destabilizing function. Mol Cell Biol **14:**416-26.

21. **Chen, F., C. MacDonald, and J. Wilusz.** 1995. Cleavage site determinants in the mammalian polyadenylation signal. Nucleic Acids Research **23:**2614-2620.

22. **Chen-Bettecken, U., E. Wecker, and A. Schimpl.** 1987. Transcriptional control of mu- and kappa-gene expression in resting and bacterial lipopolysaccharide-activated normal B cells. Immunobiology **174:**162-76.

23. **Cheng, C., and P. A. Sharp.** 2003. RNA polymerase II accumulation in the promoter-proximal region of the dihydrofolate reductase and gamma-actin genes. Mol Cell Biol **23:**1961-7.

24. **Chennathukuzhi, V., S. LeFrancois, C. Morales, V. Syed, and N. Hecht.** 2001. Elevated levels of the polyadenylation factor CstF-64 enhance formation of the 1kB testis brain RNA-binding protein (TB-RBP) mRNA in male germ cells. Mol Reprod & Dev **58:**460-469.

25. **Chuvpilo, S., M. Zimmer, A. Kerstan, J. Glockner, A. Avots, C. Escher, C. Fischer, I. Inashkina, E. Jankevics, F. Berberich-Siebelt, E. Schmitt, and E. Serfling.** 1999. Alternative polyadenylation events contribute to the induction of NF-ATc in effector T cells. Immunity **10:**261-9.

26. **Cok, S. J., S. J. Acton, A. E. Sexton, and A. R. Morrison.** 2004. Identification of RNA-binding Proteins in RAW 264.7 Cells That Recognize a Lipopolysaccharide-responsive Element in the 3-Untranslated Region of the Murine Cyclooxygenase-2 mRNA. J Biol Chem **279:**8196-205. Epub 2003 Dec 8.

27. **Cok, S. J., and A. R. Morrison.** 2001. The 3'-untranslated region of murine cyclooxygenase-2 contains multiple regulatory elements that alter message stability and translational efficiency. J Biol Chem **276:**23179-85.

28. **Cooper, C., D. Johnson, C. Roman, N. Avitahl, P. Tucker, and K. Calame.** 1992. The C/EBP family of transcriptional activators is functionally important for Ig VH promoter activity in vivo and in vitro. J Immunol **149:**3225-31.

29. **Corden, J. L.** 1990. Tails of RNA polymerase II. Trends Biochem Sci **15:**383-7.

30. **Cosma, M. P.** 2002. Ordered recruitment: gene-specific mechanism of transcription activation. Mol Cell **10:**227-36.

31. **Crawford, E. K., J. E. Ensor, I. Kalvakolanu, and J. D. Hasday.** 1997. The role of 3' poly(A) tail metabolism in tumor necrosis factor-alpha regulation. J Biol Chem **272:**21120-7.

32. **Dantonel, J.-C., K. Murthy, J. Manley, and L. Tora.** 1997. Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA. Nature **389:**399-402.

33. **Dass, B., E. N. Attaya, A. M. Wallace, and C. C. MacDonald.** 2001. Overexpression of the CstF-64 and CPSF-160 polyadenylation protein messenger RNAs in mouse male germ cells. Biol of Reproduction **64:**1722-9.

34. **Dass, B., L. McDaniel, R. A. Schultz, E. Attaya, and C. C. MacDonald.** 2002. The gene CSTF2T, encoding the human variant CstF-64 polyadenylation protein tauCstF-64, lacks introns and may be associated with male sterility. Genomics **80:**509-14.

35. **Dass, B., W. McMahon, N. A. Jenkins, D. J. Gilbert, N. G. Copeland, and C. MacDonald.** 2001. The gene for a variant form of the polyadenylation protein CstF-64 is on chromosome 19 and is expressed in pachytene spermatocyctes in mice. J Biol Chem **11:**8044-8050.

36. **DeMaria, C. T., and G. Brewer.** 1996. AUF1 binding affinity to A+U-rich elements correlates with rapid mRNA degradation. J Biol Chem **271:**12179-84.

37. **Dichtl, B., D. Blank, M. Sadowski, W. Hubner, S. Weiser, and W. Keller.** 2002. Yhh1p/Cft1p directly links poly(A) site recognition and RNA polymerase II transcription termination. EMBO J. **21:**4125-4135.

38. **Dixon, D. A., C. D. Kaplan, T. M. McIntyre, G. A. Zimmerman, and S. M. Prescott.** 2000. Post-transcriptional control of cyclooxygenase-2 gene expression. The role of the 3'-untranslated region. J Biol Chem **275:**11750-7.

39. **Doenecke, A., E. L. Winnacker, and M. Hallek.** 1997. Rapid amplification of cDNA ends (RACE) improves the PCR-based isolation of immunoglobulin variable region genes from murine and human lymphoma cells and cell lines. Leukemia **11:**1787-92.

40. **Doherty, J. K., C. T. Bond, W. Hua, J. P. Adelman, and G. M. Clinton.** 1999. An alternative HER-2/neu transcript of 8 kb has an extended 3'UTR and displays increased stability in SKOV-3 ovarian carcinoma cells. Gynecol Oncol **74:**408-15.

41. **DuBois, R. N., M. Tsujii, P. Bishop, J. A. Awad, K. Makita, and A. Lanahan.** 1994. Cloning and characterization of a growth factor-inducible cyclooxygenase gene from rat intestinal epithelial cells. Am J Physiol **266:**G822-7.

42. **Eaton, S., and K. Calame.** 1987. Multiple DNA sequence elements are necessary for the function of an immunoglobulin heavy chain promoter. Proc Natl Acad Sci U S A **84:**7634-8.

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

44. **Edwalds-Gilbert, G., and C. Milcarek.** 1995. Regulation of poly(A) site use during mouse B-cell development involves a change in the binding of a general polyadenylation factor in a B-cell stage-specific manner. Molecular and Cellular Biology **15:**6420-6429.

45. **Edwalds-Gilbert, G., K. L. Veraldi, and C. Milcarek.** 1997. Alternative poly(A) site selection in complex transcription units: means to an end? Nucleic Acids Research **25:**2547-61.

46. **Fan, X. C., and J. A. Steitz.** 1998. Overexpression of HuR, a nuclear-cytoplasmic shuttling protein, increases the in vivo stability of ARE-containing mRNAs. Embo J **17:**3448-60.

47. **Fong, N., and D. L. Bentley.** 2001. Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. Genes Dev **15:**1783-95.

48. **Fong, N., G. Bird, M. Vigneron, and D. L. Bentley.** 2003. A 10 residue motif at the C-terminus of the RNA pol II CTD is required for transcription, splicing and 3' end processing. Embo J **22:**4274-82.

49. **Foss, D. L., M. J. Baarsch, and M. P. Murtaugh.** 1998. Regulation of hypoxanthine phosphoribosyltransferase, glyceraldehyde-3-phosphate dehydrogenase and beta-actin mRNA expression in porcine immune cells and tissues. Anim Biotechnol **9:**67-78.

50. **Gamberi, C., E. Izaurralde, C. Beisel, and I. Mattaj.** 1997. Interaction between the human nuclear cap-binding protein complex and hnRNP F. Molecular & Cellular Biology **17:**2587-97.

51. **Ge, H., and R. G. Roeder.** 1994. Purification, cloning, and characterization of a human coactivator, PC4, that mediates transcriptional activation of class II genes. Cell **78:**513-523.

52. **Gehring, N. H., U. Frede, G. Neu-Yilik, P. Hundsdoerfer, B. Vetter, M. W. Hentze, and A. E. Kulozik.** 2001. Increased efficiency of mRNA 3' end formation: a new genetic mechanism contributing to hereditary thrombophilia. Nat Genet **28:**389-92.

53. **Gerber, W. V., S. A. Vokes, N. R. Zearfoss, and P. A. Krieg.** 2002. A role for the RNA-binding protein, hermes, in the regulation of heart development. Dev Biol **247:**116-26.

54. **Goethe, R., and L. Phi-van.** 1998. Posttranscriptional lipopolysaccharide regulation of the lysozyme gene at processing of the primary transcript in myelomonocytic HD11 cells. J Immunol **160:**4970-8.

55. **Greenbaum, S., and Y. Zhuang.** 2002. Identification of E2A target genes in B lymphocyte development by using a gene tagging-based chromatin immunoprecipitation system. Proc Natl Acad Sci U S A **99:**15030-5. Epub 2002 Nov 1.

56. **Grolleau, A., D. E. Misek, R. Kuick, S. Hanash, and J. J. Mule.** 2003. Inducible expression of macrophage receptor Marco by dendritic cells following phagocytic uptake of dead cells uncovered by oligonucleotide arrays. J Immunol **171:**2879-88.

57. **Gross, S., and C. L. Moore.** 2001. Rna15 Interaction with the A-Rich Yeast Polyadenylation Signal Is an Essential Step in mRNA 3'-End Formation. Mol. Cell. Biol. **21:**8045-8055.

58. **Grosschedl, R., and D. Baltimore.** 1985. Cell-type specificity of immunoglobulin gene expression is regulated by at least three DNA sequence elements. Cell **41:**885-97.

59. **Gueydan, C., L. Droogmans, P. Chalon, G. Huez, D. Caput, and V. Kruys.** 1999. Identification of TIAR as a protein binding to the translational regulatory AU-rich element of tumor necrosis factor alpha mRNA. J Biol Chem **274:**2322-6.

60. **Handel, M. A., P. A. Hunt, M. C. Kot, C. Park, and M. Shannon.** 1991. Role of sex chromosomes in the control of male germ-cell differentiation. Ann N Y Acad Sci **637:**64-73.

61. **Hatton, L. S., J. J. Eloranta, L. M. Figueiredo, Y. Takagaki, J. L. Manley, and K. O'Hare.** 2000. The Drosophila homologue of the 64 kDa subunit of cleavage stimulation factor interacts with the 77 kDa subunit encoded by the suppressor of forked gene. Nucleic Acids Res **28:**520-6.

62. **Hauksdottir, H., B. Farboud, and M. L. Privalsky.** 2003. Retinoic acid receptors beta and gamma do not repress, but instead activate target gene transcription in both the absence and presence of hormone ligand. Mol Endocrinol **17:**373-85. Epub 2002 Dec 23.

63. **Hedley, M. L., and T. Maniatis.** 1991. Sex-specific splicing and polyadenylation of dsx pre-mRNA requires a sequence that binds specifically to tra-2 protein in vitro. Cell **65:**579-86.

64. **Higgs, D. R., S. E. Goodbourn, J. Lamb, J. B. Clegg, D. J. Weatherall, and N. J. Proudfoot.** 1983. Alpha-thalassaemia caused by a polyadenylation signal mutation. Nature **306:**398-400.

65. **Hirose, Y., and J. L. Manley.** 1998. RNA polymerase II is an essential mRNA polyadenylation factor. Nature **395:**93-6.

66. **Hofmann, I., M. Schnolzer, I. Kaufmann, and W. W. Franke.** 2002. Symplekin, a constitutive protein of karyo- and cytoplasmic particles involved in mRNA biogenesis in Xenopus laevis oocytes. Mol Biol Cell **13:**1665-76.

67. **Honore, B., H. Rasmussen, H. Vorum, K. Dejgaard, X. Liu, P. Gromov, P. Madsen, B. Gesser, N. Tommerup, and J. Celis.** 1995. Heterogeneous nuclear ribonucleoproteins H, H', and F are members of a ubiquitously expressed subfamily of related but distinct proteins encoded by genes mapping to different chromomsomes. Journal of Biological Chemistry **270:**28780-28789.

68. **Honore, B., H. Vorum, and U. Baandrup.** 1999. hnRNPs H, H' and F behave differently with respect to posttranslational cleavage and subcellular localization. FEBS Lett **456:**274-280.

69. **Jackson, D. A., A. Pombo, and F. Iborra.** 2000. The balance sheet for transcription: an analysis of nuclear RNA metabolism in mammalian cells. Faseb J **14:**242-54.

70. **Kadener, S., J. P. Fededa, M. Rosbash, and A. R. Kornblihtt.** 2002. Regulation of alternative splicing by a transcriptional enhancer through RNA pol II elongation. Proc Natl Acad Sci U S A **99:**8185-90.

71. **Keller, W., S. Bienroth, K. M. Lang, G. Christofori, E. Wahle, and C. Suter-Crazzolara.** 1991. Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3' processing signal AAUAAA. Purification of the cleavage and polyadenylation factor involved in the 3'-processing of messenger RNA precursors. Embo J **10:**4241-9.

72. **Kepka-Lenhart, D., S. K. Mistry, G. Wu, and S. M. Morris, Jr.** 2000. Arginase I: a limiting factor for nitric oxide and polyamine synthesis by activated macrophages? Am J Physiol Regul Integr Comp Physiol **279:**R2237-42.

73. **Kim, M., S. H. Ahn, N. J. Krogan, J. F. Greenblatt, and S. Buratowski.** 2004. Transitions in RNA polymerase II elongation complexes at the 3' ends of genes. Embo J **23:**354-64. Epub 2004 Jan 22.

74. **Kleiman, F., and J. Manley.** 1999. Functional interaction of BRCA1-associated BARD1 with polyadenylation factor CstF-50. Science **285:**1576-9.

75. **Komarnitsky, P., E. J. Cho, and S. Buratowski.** 2000. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev **14:**2452-60.

76. **Kontoyiannis, D., M. Pasparakis, T. T. Pizarro, F. Cominelli, and G. Kollias.** 1999. Impaired on/off regulation of TNF biosynthesis in mice lacking TNF AU-rich elements: implications for joint and gut-associated immunopathologies. Immunity **10:**387-98.

77. **Kyburz, A., M. Sadowski, B. Dichtl, and W. Keller.** 2003. The role of the yeast cleavage and polyadenylation factor subunit Ydh1p/Cft2p in pre-mRNA 3'-end formation. Nucleic Acids Res **31:**3936-45.

78. **Lazarov, M. E., M. M. Martin, B. M. Willardson, and T. S. Elton.** 1999. Human phosducin-like protein (hPhLP) messenger RNA stability is regulated by cis-acting instability elements present in the 3'-untranslated region. Biochim Biophys Acta **1446:**253-64.

79. **Legendre, M., and D. Gautheret.** 2003. Sequence determinants in human polyadenylation site selection. BMC Genomics **4:**7.

80. **Lei, E. P., H. Krebber, and P. A. Silver.** 2001. Messenger RNAs are recruited for nuclear export during transcription. Genes Dev **15:**1771-82.

81. **Licatalosi, D. D., G. Geiger, M. Minet, S. Schroeder, K. Cilli, J. B. McNeil, and D. L. Bentley.** 2002. Functional interaction of yeast pre-mRNA 3' end processing factors with RNA polymerase II. Mol Cell **9:**1101-11.

82. **Macknight, R., I. Bancroft, T. Page, C. Lister, R. Schmidt, K. Love, L. Westphal, G. Murphy, S. Sherson, C. Cobbett, and C. Dean.** 1997. FCA, a gene controlling flowering time in Arabidopsis, encodes a protein containing RNA-binding domains. Cell **89:**737-45.

83. **Majello, B., and G. Napolitano.** 2001. Control of RNA polymerase II activity by dedicated CTD kinases and phosphatases. Front Biosci **6:**D1358-68.

84. **Mann, K. P., E. A. Weiss, and J. R. Nevins.** 1993. Alternative Poly(A) Site Utilization during Adenovirus Infection Coincides with a Decrease in the Activity of a Poly(A) Site Processing Factor. Molecular and Cellular Biology **13:**2411-2419.

85. **Marshall, N. F., J. Peng, Z. Xie, and D. H. Price.** 1996. Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. J Biol Chem **271:**27176-83.

86. **Martincic, K., R. Campbell, G. Edwalds-Gilbert, L. Souan, M. T. Lotze, and C. Milcarek.** 1998. Increase in the 64-kDa subunit of the polyadenylation/cleavage stimulatory factor during the G0 to S phase transition. Proceedings of the National Academy of Sciences of the United States of America **95:**11095-100.

87. **Matis, S. A., K. Martincic, and C. Milcarek.** 1996. B-lineage regulated polyadenylation occurs on weak poly(A) sites regardless of sequence composition at the cleavage and downstream regions. Nucleic Acids Research **24:**4684-92.

88. **Mbella, E. G., S. Bertrand, G. Huez, and J. N. Octave.** 2000. A GG nucleotide sequence of the 3' untranslated region of amyloid precursor protein mRNA plays a key role in the regulation of translation and the binding of proteins. Mol Cell Biol **20:**4572-9.

89. **McCracken, S., N. Fong, K. Yankulov, S. Ballantyne, G. Pan, J. Greenblatt, S. D. Patterson, M. Wickens, and D. L. Bentley.** 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature **385:**357-361.

90. **McGregor, F., A. Phelan, J. Dunlop, and J. B. Clements.** 1996. Regulation of herpes simplex virus poly(A) site usage and the action of immediate-early protein IE63 in the early-late switch. Journal of Virology **70:**1931-1940.

91. **Meinhart, A., and P. Cramer.** 2004. Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. Nature **430:**223-6.

92. **Milcarek, C., and B. Hall.** 1985. Cell-specific expression of secreted versus membrane forms of immunoglobulin gamma 2b mRNA involves selective use of alternate polyadenylation sites. Molecular & Cellular Biology **5:**2514-20.

93. **Milcarek, C., M. Suda-Hartman, and S. C. Croll.** 1996. Changes in abundance of IgG 2a mRNA in the nucleus and cytoplasm of a murine B-lymphoma before and after fusion to a myeloma cell. Molecular Immunology **33:**691-701.

94. **Mitchelson, A., M. Simonelig, C. Williams, and K. O'Hare.** 1993. Homology with Saccharomyces cerevisiae RNA14 suggests that phenotypic suppression in Drosophila melanogaster by suppressor of forked occurs at the level of RNA stability. Genes Dev **7:**241-9.

95. **Mulligan, R. C., and P. Berg.** 1980. Expression of a bacterial gene in mammalian cells. Science **209:**1422-7.

96. **Murthy, K., and J. Manley.** 1992. Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. Journal of Biological Chemistry **267:**14804-14811.

97. **Myer, V. E., X. C. Fan, and J. A. Steitz.** 1997. Identification of HuR as a protein implicated in AUUUA-mediated mRNA decay. Embo J **16:**2130-9.

98. **Natalizio, B. J., L. C. Muniz, G. K. Arhin, J. Wilusz, and C. S. Lutz.** 2002. Upstream elements present in the 3'-untranslated region of collagen genes influence the processing efficiency of overlapping polyadenylation signals. J Biol Chem **277:**42733-40. Epub 2002 Aug 27.

99. **Navalgund, L. G., C. Rossana, A. J. Muench, and L. F. Johnson.** 1980. Cell cycle regulation of thymidylate synthetase gene expression in cultured mouse fibroblasts. J Biol Chem **255:**7386-90.

100. **Newton, R., J. Seybold, L. M. Kuitert, M. Bergmann, and P. J. Barnes.** 1998. Repression of cyclooxygenase-2 and prostaglandin E2 release by dexamethasone occurs by transcriptional and post-transcriptional mechanisms involving loss of polyadenylated mRNA. J Biol Chem **273:**32312-21.

101. **Newton, R., J. Seybold, S. F. Liu, and P. J. Barnes.** 1997. Alternate COX-2 transcripts are differentially regulated: implications for post-transcriptional control. Biochem Biophys Res Commun **234:**85-9.

102. **O'Banion, M. K., V. D. Winn, and D. A. Young.** 1992. cDNA cloning and functional activity of a glucocorticoid-regulated inflammatory cyclooxygenase. Proc Natl Acad Sci U S A **89:**4888-92.

103. **Oelgeschlager, T.** 2002. Regulation of RNA polymerase II activity by CTD phosphorylation and cell cycle control. J Cell Physiol **190:**160-9.

104. **O'Neill, L. A.** 2002. Toll-like receptor signal transduction and the tailoring of innate immunity: a role for Mal? Trends Immunol **23:**296-300.

105. **Palancade, B., and O. Bensaude.** 2003. Investigating RNA polymerase II carboxyl-terminal domain (CTD) phosphorylation. Eur J Biochem **270:**3859-70.

106. **Pauws, E., A. H. van Kampen, S. A. van de Graaf, J. J. de Vijlder, and C. Ris-Stalpers.** 2001. Heterogeneity in polyadenylation cleavage sites in mammalian mRNA sequences: implications for SAGE analysis. Nucleic Acids Res **29:**1690-4.

107. **Peng, S. S., C. Y. Chen, and A. B. Shyu.** 1996. Functional characterization of a non-AUUUA AU-rich element from the c-jun proto-oncogene mRNA: evidence for a novel class of AU-rich elements. Mol Cell Biol **16:**1490-9.

108. **Perez Canadillas, J. M., and G. Varani.** 2003. Recognition of GU-rich polyadenylation regulatory elements by human CstF-64 protein. Embo J **22:**2821-30.

109. **Peterson, M. L., S. Bertolino, and F. Davis.** 2002. An RNA polymerase pause site is associated with the immunoglobulin mus poly(A) site. Mol Cell Biol **22:**5606-15.

110. **Peterson, M. L., E. R. Gimmi, and R. P. Perry.** 1991. The developmentally regulated shift from membrane to secreted mu mRNA production is accompanied by an increase in cleavage-polyadenylation efficiency but no measurable change in splicing efficiency. Molecular and Cellular Biology **11:**2324-2327.

111. **Piecyk, M., S. Wax, A. R. Beck, N. Kedersha, M. Gupta, B. Maritim, S. Chen, C. Gueydan, V. Kruys, M. Streuli, and P. Anderson.** 2000. TIA-1 is a translational silencer that selectively regulates the expression of TNF-alpha. Embo J **19:**4154-63.

112. **Poellinger, L., B. K. Yoza, and R. G. Roeder.** 1989. Functional cooperativity between protein molecules bound at two distinct sequence elements of the immunoglobulin heavy-chain promoter. Nature **337:**573-6.

113. **Pokholok, D. K., N. M. Hannett, and R. A. Young.** 2002. Exchange of RNA polymerase II initiation and elongation factors during gene expression in vivo. Mol Cell **9:**799-809.

114. **Prelich, G.** 2002. RNA polymerase II carboxy-terminal domain kinases: emerging clues to their function. Eukaryot Cell **1:**153-62.

115. **Prendergast, G. C., and E. B. Ziff.** 1989. DNA-binding motif. Nature **341:**392.

116. **Quesada, V., R. Macknight, C. Dean, and G. G. Simpson.** 2003. Autoregulation of FCA pre-mRNA processing controls Arabidopsis flowering time. Embo J **22:**3142-52.

117. **Quintans, J., and I. Lefkovits.** 1974. Clonal expansion of lipopolysaccharide-stimulated B lymphocytes. J Immunol **113:**1373-6.

118. **Ramanathan, Y., S. M. Rajpara, S. M. Reza, E. Lees, S. Shuman, M. B. Mathews, and T. Pe'ery.** 2001. Three RNA polymerase II carboxyl-terminal domain kinases display distinct substrate preferences. J Biol Chem **276:**10913-20. Epub 2001 Jan 16.

119. **Rickert, P., J. L. Corden, and E. Lees.** 1999. Cyclin C/CDK8 and cyclin H/CDK7/p36 are biochemically distinct CTD kinases. Oncogene **18:**1093-102.

120. **Rimokh, R., F. Berger, C. Bastard, B. Klein, M. French, E. Archimbaud, J. P. Rouault, B. Santa Lucia, L. Duret, M. Vuillaume, and et al.** 1994. Rearrangement of CCND1 (BCL1/PRAD1) 3' untranslated region in mantle-cell lymphomas and t(11q13)-associated leukemias. Blood **83:**3689-96.

121. **Ron, D., and J. F. Habener.** 1992. CHOP, a novel developmentally regulated nuclear protein that dimerizes with transcription factors C/EBP and LAP and functions as a dominant-negative inhibitor of gene transcription. Genes Dev **6:**439-53.

122. **Rossignol, M., I. Kolb-Cheynel, and J. M. Egly.** 1997. Substrate specificity of the cdk-activating kinase (CAK) is altered upon association with TFIIH. Embo J **16:**1628-37.

123. **Ryan, K., O. Calvo, and J. L. Manley.** 2004. Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease. Rna **10:**565-73.

124. **Ryan, K., K. G. Murthy, S. Kaneko, and J. L. Manley.** 2002. Requirements of the RNA polymerase II C-terminal domain for reconstituting pre-mRNA 3' cleavage. Mol Cell Biol **22:**1684-92.

125. **Schroeder, S. C., B. Schwer, S. Shuman, and D. Bentley.** 2000. Dynamic association of capping enzymes with transcribing RNA polymerase II. Genes Dev **14:**2435-40.

126. **Schul, W., B. Groenhout, K. Koberna, Y. Takagaki, A. Jenny, E. M. M. Manders, I. Raska, R. vanDriel, and L. deJong.** 1996. The RNA 3' cleavage factors CstF-64 kDa and CPSF-100 kDa are concentrated in nuclear domains closely associated with coiled bodies and newly synthesized RNA. EMBO Journal **15:**2883-2892.

127. **Schul, W., I. v. d. Kraan, A. G. Matera, and R. v. Driel.** 1999. Nuclear domains enriched in RNA 3'-processing factors associate with coiled bodies and histone genes in a cell cycle-dependent manner. Mol Biol Cell **10:**3815-3824.

128. **Schwartz, L. B., and R. G. Roeder.** 1975. Purification and subunit structure of deoxyribonucleic acid-dependent ribonucleic acid polymerase II from the mouse plasmacytoma, MOPC 315. J. Biol. Chem. **250:**3221-3228.

129. **Sheets, M. D., S. C. Ogg, and M. P. Wickens.** 1990. Point mutations in AAUAAA and the poly(A) addition site: effects on the accuracy and effiiciency of cleavage and polyadenylation in vitro. Nucleic Acids Research **18:**5799-5805.

130. **Shilatifard, A.** 1998. Factors regulating the transcriptional elongation activity of RNA polymerase II. FASEB J. **12:**1437-1446.

131. **Shilatifard, A., D. R. Duan, D. Haque, C. Florence, W. H. Schubach, J. W. Conaway, and R. C. Conaway.** 1997. ELL2, a new member of an ELL family of RNA polymerase II eleongation factors. Proc. Natl. Acad. Sci. USA **94:**3639-3643.

132. **Simpson, G. G., P. P. Dijkwel, V. Quesada, I. Henderson, and C. Dean.** 2003. FY is an RNA 3' end-processing factor that interacts with FCA to control the Arabidopsis floral transition. Cell **113:**777-87.

133. **Sun, X. H., N. G. Copeland, N. A. Jenkins, and D. Baltimore.** 1991. Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins. Mol Cell Biol **11:**5603-11.

134. **Takagaki, Y., C. C. MacDonald, T. Shenk, and J. L. Manley.** 1992. The human 64-kDa polyadenylation factor contains a ribonucleoprotein-type RNA binding domain and unusual auxilary motifs. Proc Natl Acad Sci USA **89:**1403-1407.

135. **Takagaki, Y., and J. Manley.** 1994. A polyadenylation factor subunit is the human homologue of the *Drosophila suppressor of forked* protein. Nature **372:**471-474.

136. **Takagaki, Y., and J. L. Manley.** 2000. Complex protein interactions within the human polyadenylation machinery identify a novel component. Mol Cell Biol **20:**1515-1525.

137. **Takagaki, Y., and J. L. Manley.** 1992. A human polyadenylation factor is a G protein beta-subunit homologue. Journal of Biological Chemistry **267:**23471-23474.

138. **Takagaki, Y., and J. L. Manley.** 1997. RNA recognition by the human polyadenylation factor CstF. Molecular and Cellular Biology **17:**3907-3914.

139. **Takagaki, Y., R. Seipelt, M. Peterson, and J. Manley.** 1996. The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B-cell differentiation. Cell **87:**941-952.

140. **Tian, B., J. Hu, H. Zhang, and C. S. Lutz.** 2005. A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucleic Acids Res **33:**201-12. Print 2005.

141. **Trigon, S., H. Serizawa, J. W. Conaway, R. C. Conaway, S. P. Jackson, and M. Morange.** 1998. Characterization of the residues phosphorylated in vitro by different C-terminal domain kinases. J Biol Chem **273:**6769-75.

142. **Ubeda, M., M. Vallejo, and J. F. Habener.** 1999. CHOP enhancement of gene transcription by interactions with Jun/Fos AP-1 complex proteins. Mol Cell Biol **19:**7589-99.

143. **Underhill, D. M., and A. Ozinsky.** 2002. Toll-like receptors: key mediators of microbe detection. Curr Opin Immunol **14:**103-10.

144. **Vadiveloo, P. K., G. Vairo, U. Novak, A. K. Royston, G. Whitty, E. L. Filonzi, E. J. Cragoe, Jr, and J. A. Hamilton.** 1996. Differential regulation of cell cycle machinery by various antiproliferative agents is linked to macrophage arrest at distinct G1 checkpoints. Oncogene **13:**599-608.

145. **Vairo, G., A. K. Royston, and J. A. Hamilton.** 1992. Biochemical events accompanying macrophage activation and the inhibition of colony-stimulating factor-1-induced macrophage proliferation by tumor necrosis factor-alpha, interferon-gamma, and lipopolysaccharide. Journal Of Cellular Physiology **151:**630-641.

146. **Veraldi, K. L., G. K. Arhin, K. Martincic, L. H. Chung-Ganster, J. Wilusz, and C. Milcarek.** 2001. hnRNP F influences binding of a 64-kilodalton subunit of cleavage stimulation factor to mRNA precursors in mouse B cells. Molecular & Cellular Biology **21:**1228-38.

147. **Wallace, A., B. Dass, S. Ravnik, V. Tonk, N. Jenkins, D. Gilbert, N. Copeland, and C. MacDonald.** 1999. Two distinct forms of the 64,000 Mr protein of the cleavage stimulation factor are expressed in mouse male germ cells. Proc Natl Acad Sci USA **96:**6763-6768.

148. **Weiss, E. A., A. Michael, and D. Yuan.** 1989. Role of transcriptional termination in the regulation of mu mRNA expression in B lymphocytes. J Immunol **143:**1046-52.

149. **Wiersma, E. J., D. Ronai, M. Berru, F. W. Tsui, and M. J. Shulman.** 1999. Role of the intronic elements in the endogenous immunoglobulin heavy chain locus. Either the matrix attachment regions or the core enhancer is sufficient to maintain expression. J Biol Chem **274:**4858-62.

150. **Wilusz, C. J., M. Wormington, and S. W. Peltz.** 2001. The cap-to-tail guide to mRNA turnover. Nat Rev Mol Cell Biol **2:**237-46.

151. **Wilusz, J., S. Pettine, and T. Shenk.** 1989. Functional analysis of point mutations in the AAUAAA motif of the SV40 late polyadenylation signal. Nucleic Acids Research **17:**3899-3908.

152. **Wu, C., and J. C. Alwine.** 2004. Secondary structure as a functional feature in the downstream region of mammalian polyadenylation signals. Mol Cell Biol **24:**2789-96.

153. **Yamshchikov, V. F., M. Mishina, and F. Cominelli.** 2002. A possible role of IL-1ra 3'-untranslated region in modulation of protein production. Cytokine **17:**98-107.

154. **Yankulov, K., K. Yamashita, R. Roy, J. M. Egly, and D. L. Bentley.** 1995. The transcriptional elongation inhibitor 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole inhibits transcription factor IIH-associated protein kinase. J Biol Chem **270:**23922-5.

155. **Zearfoss, N. R., A. P. Chan, C. F. Wu, M. Kloc, and L. D. Etkin.** 2004. Hermes is a localized factor regulating cleavage of vegetal blastomeres in Xenopus laevis. Dev Biol **267:**60-71.

156. **Zhang, J., and J. Corden.** 1991. Identification of phosphorylation sites in the repetitive carboxyl- terminal domain of the mouse RNA polymerase II largest subunit. J. Biol. Chem. **266:**2290-2296.

157. **Zhang, T., V. Kruys, G. Huez, and C. Gueydan.** 2002. AU-rich element-mediated translational control: complexity and multiple activities of trans-activating factors. Biochem Soc Trans **30:**952-8.

158. **Zhao, J., L. Hyman, and C. Moore.** 1999. Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. Microbiol Mol Biol Rev **63:**405-45.