Sequence variations affecting AU-rich element function and disease

Edward Hitti1, Khalid S. A. Khabar1

1BioMolecular Research Program, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. Alternative splicing and alternative polyadenylation, au-rich element function and disease
   3.1. Alternative splicing affecting ARE function
       3.1.1. The ARE-binding protein AUF1
       3.1.2. Beta-catenin
       3.1.3. The T-cell receptor zeta chain
   3.2. Alternative polyadenylation affecting ARE function
       3.2.1. The AU-rich binding protein HuR
       3.2.2. Fibroblast growth factor 2
       3.2.3. Cyclooxygenase
       3.2.4. Steroidogenic acute regulatory protein
       3.2.5. VEGF receptor 1
       3.2.6. Casein kinase Alpha
       3.2.7. Phosducin-like protein
4. Genetic polymorphisms in or near AU-rich elements
   4.1. Single nucleotide polymorphism in or near AU-rich elements
       4.1.1. Bone morphogenetic protein 2
       4.1.2. RET
       4.1.3. Thrombin-activable fibrinolysis inhibitor
       4.1.4. Human glycoprotein PC-1
       4.1.5. Human glucocorticoid receptor
   4.2. Deletions/insertions AU-rich element function and disease
       4.2.1. PPP1R3 a regulator of muscle glycogen
       4.2.2. Beta 2-adrenergic receptor
       4.2.3. Low-density lipoprotein receptor
       4.2.4. Thymidylate synthase
5. Perspectives
6. References

1. ABSTRACT

Adenylate-uridylate rich elements (AREs) in the 3'UTRs of many transiently expressed genes regulate mRNA instability and translation. Such ARE-genes are involved in vital biological processes like cellular growth, differentiation, and immunity. Defects in their expression contribute to a variety of disease conditions like cancer, autoimmune diseases, diabetes, and cardiovascular and chronic inflammatory diseases. Over the past two decades, considerable progress has been made in understanding the mode of regulation of AREs containing mRNAs by RNA-binding proteins, miRNAs, and signaling pathways. This review focuses on the less documented sequence variation affecting ARE functions and its relation to disease. We discuss reports describing genetic polymorphisms, alternative polyadenylation, and alternative splicing that can lead to the loss or gain of function of AREs, often with significant implications to disease.

2. INTRODUCTION

The regulation of gene expression at the post-transcriptional level is vital for normal cell function, and aberrations contribute to a wide variety of diseases and disorders (1-2). During their biogenesis, pre-mRNAs associate with a host of RNA-binding proteins (RBPs) that mediate diverse RNA processing reactions including 5'-end capping, splicing, editing, 3'-end cleavage, polyadenylation, and transport (Reviewed in:(1)). The transcripts are subsequently exported through nuclear pores to the cytoplasm (3). The cytoplasmic regulation of the stability and translation of mRNAs are fine-tuned steps that provide the ability to respond early to external stimuli and ensures specificity of the post-transcriptional regulatory process. This specificity is achieved by sequence regulatory elements that can be present in any region in the transcript including the coding region and the 5' and 3' untranslated regions (UTRs) (4-5). One group of regulatory sequences
are the AU-rich elements (AREs), a heterogeneous class of adenylate uridylate-rich sequences in the 3'UTR on many mRNAs. Chen and Shyu grouped AREs in three classes. In class I AUUUA motifs are scattered and coupled with nearby U-rich regions or U stretches. Class II AREs have at least two overlapping copies of the motif UAUAAA(U/A)(U/A) in a U-rich region; class III AREs do not contain any AUUUA motifs but are overall rich in Us and As. All three types can lead to strong destabilization of the mRNA (6). Functional analysis considers the nonamer UAUAUAU or its variation UAUAUAUWW as a minimum necessary for destabilization (7). A single AUUUA pentamer in a non-AU context is very common and not functional, and therefore, it cannot be considered an ARE on its own. Using a computational biology approach, we clustered human AREs in five groups considering the 13mer WWWUAUUUAUWWW as a minimal ARE; clustering was based on the number of overlapping AUUUA motifs (8). According to the most recent computational screen, AU-rich mRNAs constitute a large class of the transcriptome of up to 15% (9). The AREs of specific genes are highly conserved in evolution and their gene products have diverse functions including inflammatory and immune response, transcription, cellular proliferation, RNA metabolism, development, and signaling (9-10). AREs can be targeted by transacting factors such as RNA-binding-proteins that modulate the stability and translation of mRNAs. Such proteins include TTP, HuR, AUF1, KSRP, TIA1, TIAR, BRF1, and BRF2 (11-12). Those proteins are themselves often targets of signaling modules that regulate their function either by phosphorylation or by the regulation of cellular localization. The main signaling pathway that regulates ARE function is the stress activated MAP kinase, p38/MK2/TTP pathway (13-14). Other signaling pathways have been implicated such as the ERK, the PI3-AKT, the PKC and the Wnt/Beta-catenin signaling pathways (15). ARE function can also be regulated by miRNAs (16-17).

The mechanisms of ARE function and regulation are well-understood; aberrations in this regulation can be major players in a variety of diseases like diabetes, cancer, and cardiovascular and inflammatory diseases. This aspect of AU-rich function is well documented and several recent reviews, including one by the current authors (2), have been published. Here the focus is on sequence variations arising from alternative splicing and alternative polyadenylation and genetic polymorphism that can affect AU-rich sequence and function in disease.

3. ALTERNATIVE SPLICING AND ALTERNATIVE POLYADENYLATION, AU-RICH ELEMENT FUNCTION AND DISEASE

Alternative splicing is a key contributor to the complexity of mammalian gene expression; it is thought that almost all human genes can be alternatively spliced (18). However, alternative splicing events within the 3'UTR of mRNAs are rare, since transcripts with such 3'UTR are usually directed for nonsense-mediated decay (NMD) (19). Therefore, only few mRNAs are known to be alternatively spliced within their 3'UTRs in physiologic situations and they include the ARE-binding protein AUF1, the splicing factor SC35, and the transcriptional coactivator Beta-catenin (20-22).

Alternative polyadenylation is also a common process that contributes to the complexity in the regulation of expression of the human transcriptome. It is believed, that in normal human cells, more than half of pre-mRNAs are alternatively polyadenylated (23). Alternative polyadenylation may cause the deletion of cis acting sequence elements resulting in differential responses to RNA-binding proteins, miRNA, or signaling events.

Aberrant splicing and aberrant polyadenylation can also be found in certain disease states. This section presents genes that are alternatively spliced or polyadenylated in regions that affect ARE presence and function. Cyclin D1 transcript can be both alternatively spliced and alternatively polyadenylated in mantle cell lymphoma. The other genes will be grouped in alternative splicing or alternative polyadenylation sections.

Cyclin D1 is an important regulator of the mitotic cell cycle. It is a proto-oncogene and its expression is highly controlled at the transcriptional, post-transcriptional, and post-translational levels. Over-expression of cyclin D1 has been linked to the development and the progression of several cancers (24-25). The gene of cyclin D1 has five exons which can be alternatively spliced to create two major mRNA isoforms, cyclin D1a and D1b. The cyclin D1a isoform is a transcript of 4.5 kb in length, with a small coding region and a very large (>3 kb) 3'UTR. The 3'UTR is overall rich in As and Us and contains seven widely spaced AUUUUA pentamers, all of which are near or within A- and U-rich contexts. The alternatively spliced cyclin D1b isoform lacks exon 5 but retains intron 4, resulting in a proximal polyadenylation signal that leads to the truncation of nearly all of its 3'UTR, including all AUUUUA pentamers (26-28). In patients with mantle cell lymphoma (MCL), a predominant and aberrantly highly expressed cyclin D transcript lacking the 3'UTR is the result of alternative polyadenylation of transcript Cyclin D1a rather than the alternatively spliced Cyclin D1b. This alternative polyadenylation is caused by a point mutation. The same point mutation and the resulting alternative polyadenylation is found in the Z-138 MCL cell line, where the half-life of the short cyclin D1a mRNA is much longer than that of the full-length mRNA (29). MDA-MB-453, a breast cancer cell line which expresses the short truncated mRNA version was resistant to Prostaglandin A 2 (30). Unexpectedly, a recent report suggests that the AUUUUA motifs in Cyclin D1a lead to the stabilization, rather than the destabilization, of a
Sequence variations in AU-rich elements and disease

reporter transcript. The destabilization is the result of the activity of microRNAs of the miR-15/16 family and the miR17-92 cluster that targets only the full length 3'UTR and fails to regulate the alternatively spliced and polyadenylated forms found in cancer (31). In conclusion, three forms of cyclin D1 mRNA have been reported: the wild type form (cyclin D1a) with full length containing 3'UTR and two 3'UTR truncated forms which result from alternative splicing or polyadenylation. Both truncated isoforms are more stable than the wild type and are over-represented and over expressed in tumors such as MCL and breast carcinoma.

3.1. Alternative splicing affecting ARE function

3.1.1. The ARE binding protein AUF1

AUF1 is a trans acting, ARE-binding protein that enhances mRNA decay in a number of genes including proto-oncogenes, growth factors, cytokines, and cell cycle-regulatory genes (32-41). AUF1 knockout mice mRNA levels and reporter protein expression (21). and one AUUUA pentamer. Gel shift assays demonstrate regions are located within intron 9. The regions are overall 9. Three closely spaced and highly conserved AU-rich regions are located within intron 9. The gene consists of ten exons and the translational termination codon is located within exon 8. Exons 9 and 10 encode exclusively 3'UTR sequences. The 3'UTR of the AUF1 transcript can be spliced in four different forms in K562 erythroleukemia cells, two of which retain intron 9. Three closely spaced and highly conserved AU-rich regions are located within intron 9. The regions are overall rich in As and Us and contain five AUUUA elements and one AUUUA pentamer. Gel shift assays demonstrate that AUF1 can bind the AU-rich regions and reduce mRNA levels and reporter protein expression (21).

Although there are no reports that directly link the ARE-affecting alternative splicing of AUF1 to disease, the potential alternative splicing-dependent variations in cellular AUF1 levels may indirectly influence the expression of other ARE-containing genes and affect the wide spectrum of their activities and associated disorders.

3.1.2. Beta-catenin

The transcriptional coactivator Beta-catenin is part of a signaling pathway that responds to the Wnt family of secreted glycoproteins and controls embryonic development and adult homeostasis. Aberrations in Wnt/Beta-catenin signaling can contribute to human diseases including congenital malformations, cancer, and osteoporosis (43). Beta-catenin/wnt signaling can also regulate gene expression at the post-transcriptional level in an ARE-dependent manner as in the case of the mRNAs of PITX2, c-Jun, cyclin D1, and cyclin D2 (44). The Beta-catenin mRNA itself contains several U-rich regions spread over the entire sequence and can be classified as class III AREs according to Chen and Shyu (6). The stability of Beta-catenin mRNA can be regulated by the ARE-binding proteins KSRP and HuR (45-47). There are three Beta-catenin mRNA splice variants that differ in their 3'UTR due to alternative splicing (22), all of which are ubiquitously expressed; however, the level of each isoform might differ depending on cell type. One report suggests that the shortest isoform is predominantly expressed during the progression of esophageal cancer (48). The AU-rich region, which is located in exon 16B, is found in each of the three transcripts and, surprisingly, it contributes to stabilization rather than destabilization of the Beta-catenin mRNA transcripts in a splicing-dependent manner. Alternative splicing and the AU-rich region inter-dependently influence Beta-catenin mRNA stability and cellular levels (22).

3.1.3. The T-cell Receptor zeta chain

The T-cell receptor zeta chain (TCR Zeta) is considered an amplification module in the TCR signaling cascade and is essential for assembly and surface expression of the TCR/CD3 complex. The TCR Zeta is down-regulated in many chronic infectious and inflammatory diseases, including systemic lupus erythematosus (SLE) (49). The wild type 3'UTR of the TCR Zeta mRNA is approximately 1 kb; however, an alternatively spliced form of 344 nt 3'UTR is predominantly found in SLE T cells compared to normal T cells. The alternatively spliced short 3'UTR TCR Zeta mRNA is unstable and weakly translated in SLE T cells (50-51). There are three AUUUA pentamers in the wild type 3'UTR of TCR Zeta mRNA; two of them are in close proximity to each other in the proximal region of the 3'UTR before the splice-deleted region and one is located inside the deleted region. Luciferase reporter assays showed that constructs containing only the proximal pentamer region from the alternative spliced form had a weaker expression than constructs containing the wild type 3'UTR. This result, together with additional mutational analysis, indicates, surprisingly, that the AUUUA containing the splice-deleted region confers stability rather than instability to the transcript (52). This might be due to the binding and stabilization by the RNA binding protein HuR (53). The aberrant, post-transcriptional regulation of TCR Zeta is an important example of alternative splicing affecting ARE function in the human disease SLE.

3.2. Alternative Polyadenylation affecting ARE function

3.2.1. The AU-rich binding protein HuR

The RNA binding protein HuR is an example where the mRNA of an ARE-binding protein contains AREs and the presence of the elements can be regulated by alternative polyadenylation. HuR is one of the most investigated ARE-binding proteins; it is ubiquitously expressed and regulates the expression of a large number of transcripts including cytokines, proto-oncogenes, apoptotic factors, cell cycle regulators, and growth factors. Aberrant activities of HuR (i.e., both over-expression and cytoplasmic compartmentalization) have been implicated in disease including many cancer types and inflammatory diseases (54-60). It is widely accepted that HuR stabilization of target mRNAs is mediated through competition with other RNA-binding proteins that destabilize ARE-mRNAs such as TTP, certain AUF1 isoforms, and KSRP (61). HuR is mainly localized within
the cell nucleus and the nucleo-cytoplasmic shuttling of HuR is believed to be a primary regulatory mechanism of HuR function (62). Recently, we have shown that multiple polyadenylation sites in HuR pre-mRNA result in the biogenesis of three HuR mRNA isoforms that differ in the size of the 3′UTR. A polyadenylation variant with distal poly (A) signal harbors functional AREs. The HuR ARE down-regulates reporter activity and is regulated by both HuR and TTP (63). It appears that a HuR auto-regulatory mechanism results from the binding of HuR to the alternative polyadenylation transcript (63). The auto-amplification of HuR may contribute further to the HuR over-expression that is seen in many cancer types.

3.2.2. Fibroblast growth factor 2

Fibroblast growth factor 2 (FGF-2) is expressed in various cell types and tissues, and has many biological roles including embryogenesis and morphogenesis, particularly in the nervous system and in bone formation. FGF-2 is a major angiogenic factor, playing a crucial role in wound healing and cardiovascular disease. It is also involved in cancer pathophysiology, notably in tumor neovascularization (64-66). FGF-2 mRNA contains a large 3′UTR which makes up to 90% of its total length (67). This 3′UTR contains eight alternative polyadenylation sites with an AU-rich destabilizing region located between the first and second polyadenylation sites. The AU-rich region is atypical since it is made of two 122-nt direct repeats; each repeat is overall rich in A and U but contains only one AUUUA pentamer. The mere presence of the AUUUA pentamer is not sufficient to destabilize FGF-2 mRNA; the presence of both tandem repeats is necessary for RNA destabilization. (67). Most interestingly, the shortest mRNA that is cleaved at the most proximal site and that is lacking the destabilizing element is predominantly present (28-100% of total FGF-2 mRNA) in three transformed cell lines (67) while it constitutes only 4.5% of the FGF-2 mRNA in primary skin fibroblasts. This suggests that post-transcriptional regulation of FGF-2 mRNA is aberrant in transformed cells due to the polyadenylation event affecting ARE presence (67).

3.2.3. Cyclooxygenase

The immediate-early cyclo-oxygenase-2 (COX-2) gene encodes an inducible prostaglandin synthase enzyme that has been implicated in inflammatory and proliferative diseases. Pharmacological inhibition of COX can provide relief from the symptoms of inflammation and pain. Non-steroidal, anti-inflammatory drugs, such as aspirin, exert their effects through inhibition of COX (68). The post-transcriptional regulation of COX-2 mRNA has been the subject of intense investigation; its 3′UTR contains a 116-nucleotide, AU-rich region made up of a cluster of six AUUUA elements, several of which overlap, that is located near the stop codon. COX-2 ARE confers rapid decay of a normally stable reporter mRNA (reviewed in (69)). In several cell types, the COX-2 gene produces two major transcripts, 4.6 kb and 2.8 kb isoforms, which are derived from alternative polyadenylation of the 3′UTR (70-71). The anti-inflammatory glucocorticoid dexamethasone potently inhibits the stabilization of COX-2 mRNA triggered by IL-1. In response to dexamethasone, the short transcript, lacking the ARE region, decays with a longer half-life than the large isoform (71). The two COX-2 transcript variants are also found in human colorectal adenocarcinoma cell lines; the short variant of COX-2 mRNA lacking the ARE is selectively stabilized upon cell growth to confluence (72).

3.2.4. Steroidogenic acute regulatory protein

Steroidogenic acute regulatory protein (StAR) mediates the intramitochondrial transport of cholesterol, a process needed for steroid biosynthesis (73). Lipoid congenital adrenal hyperplasia (LCAH), a general human steroid deficiency syndrome, results from functionally disruptive mutations within the StAR gene (74). In rat adrenal cells, the StAR mRNA is transcribed into a predominant 3.5-kb form and 1.6-kb form. The 3.5-kb form has an extended 3′UTR that contains three scattered AUUUA pentamers. This 3′UTR causes lower mRNA stability and this instability can be enhanced by treatment with Br-cAMP in an ARE-dependent fashion (75-76). Analysis of human StAR mRNA indicates the presence of two comparable mRNA isoforms but only the 3.5-kb form is subject to hormone regulation, suggesting a post-transcriptional regulatory mechanism similar to Rat (77). The ARE-binding and destabilizing protein TIS11B (BRF1) can bind the StAR mRNA 3.5-kb variant. Surprisingly, TIS11B suppression of the StAR 3.5-kb mRNA is coupled to enhanced translation leading to increased cholesterol metabolism. (78-79). Inappropriate StAR gene expression could result in effects associated with abnormalities in adrenal and gonadal function, probably including infertility (80).

3.2.5. VEGF receptor 1

Vascular endothelial growth factor (VEGF) is regarded as the most important stimulator of angiogenesis and vascular permeability. It has been the subject of extensive interest due to its role in major diseases like cancer, autoimmune disease, eye diseases, cardiovascular diseases, and others (81-84). There are two main VEGF receptors: VEGFR1 (also called fms-like tyrosine kinase-1 or Flt1), a weak receptor tyrosine kinase that is expressed in vascular endothelial cells, placental trophoblast cells, macrophages, and monocytes, and VEGFR2 (KDR1/Flk1), a potent receptor tyrosine kinase that is primarily expressed in vascular endothelial cells especially during vascular development (85). The FLT1 gene produces two proteins due to an alternative polyadenylation event. One is Flt1, a transmembrane receptor that contains 30 spliced exons and is translated into a 200-kDa protein containing an extracellular, N-terminal, ligand-binding domain, a single membrane-spanning segment, and a C-terminal intracellular segment that carries two tyrosine kinase domains. The second alternative transcript codes for soluble Flt1 (sFlt1), which shares the first 13 exons with Flt1, can be polyadenylated at three intronic sites in intron 13, and is translated into a 100-kDa protein that has the ligand-binding domain but lacks the membrane-spanning and C-terminal domains (86-89). sFlt1 is expressed by vascular endothelial cells and is thought to play a major role in regulating
Sequence variations in AU-rich elements and disease

angiogenesis by high affinity binding and antagonizing VEGF functions. sFlt1 has physiological functions in diverse tissues such as the cornea and the uterus and a pathophysiological role in pregnancy (90-91). Selective expression of sFlt1, without corresponding expression of Flt1, is necessary to maintain the avascularity of the cornea (92). Placental expression of sFlt1 increases in pregnancy, which may contribute to the development of hypertension and preeclampsia. Preeclampsia is a medical condition of hypertension and high protein levels in urine during pregnancy (93-94). According to RNAse protection assays, three sFlt1 transcript variants exist and are the result of polyadenylation at two proximal and one distal site. The 3'UTR resulting from cleavage at the distal site is 4 kb in size and contains 19 AUUUA pentamers, most of which are flanked by AU-rich sequences. It is very likely that this variant is regulated at the post-transcriptional level with possible implication in disease (89).

3.2.6. Casein kinase Alpha

Casein kinases I (CKIs) are a family of conserved, second messenger-independent, serine/threonine protein kinases found in all eukaryotes. Members of the CK1 family are involved in diverse cellular processes including membrane trafficking, circadian rhythm, cell cycle progression, chromosome segregation, apoptosis, and cellular differentiation (95). Mutations and deregulation of CK1 expression and activity have been linked to various diseases including neurodegenerative disorders such as Alzheimer's and Parkinson's disease, sleeping disorders, and cancer (96-98). Human CK1 Alpha(L) and CK1 Alpha(S) are derived from 4.2-kb and 2.4-kb mRNA transcripts generated by alternative polyadenylation. The 3'UTR in the 4.2-kb transcript is overall rich in AU and Us and contains six widely spaced AUUUA pentamers, while the 3'UTR of the 2.4-kb transcript contains only a single motif. Analysis of CK1 Alpha 3'UTR RNA sequences suggests that in HeLa cells, the longer 3'UTR transcripts degrade approximately 13 times faster than the shorter 3'UTR transcripts (99). Since expression levels and CK1 activity have been linked to several diseases, it would be interesting to investigate the levels of the alternatively spliced isoforms in disease.

3.2.7. Phosducin-like protein

Phosducin (Pd, PDCl) and phosducin-like protein (PhLP, PDCL) regulate G-protein signaling by binding G beta gamma subunits (100). G protein signaling is involved in all major disease areas such as cardiovascular, metabolic, neurodegenerative, psychiatric, cancer, and infectious diseases (101). Alternative polyadenylation sites in human PhLP lead to the production of two transcripts, 1.2 kb and 3.1 kb, respectively. The predominantly expressed long transcript contains nine AUUUA copies and two overlapping copies of the nonamer UUAUUUA(U/A)(U/A) in the 3'UTR. PhLP AREs are functional decay elements and, in vivo, the long transcript exhibits a much shorter mRNA half-life that the shorter one. Reporter assays show that the 3'UTR of the long transcript, or only the AU-rich region, confers instability to luciferase mRNA and reduces protein expression by nine-fold (102). Thus, ARE presence and function in PhLP is regulated by polyadenylation with consequences on protein expression and potential relevance in the many diseases related to G protein signaling.

4. GENETIC POLYMORPHISMS IN OR NEAR AU-RICH ELEMENTS

Genetic polymorphisms are variations in DNA sequence that exist among individuals, groups, or populations. They include single nucleotide polymorphisms (SNPs), sequence repeats, insertions, deletions, and recombinations. Polymorphisms are rarely functional; if they were, they would be called mutations and could cause disease phenotypes with high penetrance. Common polymorphisms are less likely to cause strong phenotypes. However, they can significantly contribute to disease risk through genetic and environmental interactions. SNPs are the most common form of genetic variation in the human genome. We have grouped SNPs that affect ARE function in one section and the other forms of genetic polymorphism in another section.

4.1. Single Nucleotide Polymorphism in or near AU-rich elements

4.1.1. The bone morphogenetic protein 2

Bone Morphogenetic Protein 2 (BMP2) is a member of a family of 20 proteins that are potent growth factors belonging to the transforming growth factor Beta superfamily. They are involved in processes such as embryogenesis, skeletal formation, hematopoiesis, and neurogenesis (103). The BMP2 gene has been genetically linked to disease conditions such as osteoporosis and osteoarthritis. High concentrations of BMP2 protein induce osteogenesis and low concentrations induce adipogenesis (104-105). The 3'UTR of the BMP2 gene is highly conserved and contains ten widely spaced AU-rich motifs (AUUUA or AUUUUA). One of the AUUUA elements in human BMP2 RNA is disrupted to AUUUC in an osteoporosis-associated SNP (rs15705) (106). The polymorphic element is in close proximity (six bases distance) to another AUUUA element in a region that is AU-rich. The mRNA containing the conserved AU-rich region is very labile with a half-life of less than 10 min in extracts from undifferentiated F9 cells, a cellular model of BMP2 regulation (104). The activity of a reporter fused with the 3'UTR that contains the rs15705 C allele was significantly higher than that of the A allele 3'UTR (105). Certain proteins lost their RNA-binding ability to the C allele construct, and interestingly, an unidentified 100-kDa protein gained affinity to RNA due to the ARE-disrupting SNP (12, 104). The rs15705 SNP causes differences in mesenchymal cells and in muscle/fat phenotypes in humans and is associated with different body morphologies. Males with the C/C genotype have less subcutaneous arm fat than males with the A/C and A/A genotypes. Males with the C/C genotype gained significantly more skeletal muscle after a 12-week resistance-training program (105). The rs15705 genetic
variation is one of the few examples where a SNP in ARE was experimentally implicated in susceptibility to diseases like osteoporosis and other diseases associated with bone (104-105).

4.1.2. RET proto-oncogene

The RET proto-oncogene encodes a common receptor for the glial-cell-line-derived neurotrophic factor family of extracellular signaling molecules. RET gain-of-function mutations are associated with the development of various types of human cancer (107). While RET loss of function is associated with the development of Hirschsprung's disease (108). Hirschsprung’s disease is a complex genetic congenital disorder of the colon in which the ganglia nerve cells are absent, causing chronic constipation. The RET proto-oncogene is the major gene involved in Hirschsprung’s disease pathogenesis, with classical coding mutations accounting for up to 35% of sporadic patients (109). The 3’UTR of the mRNA of RET contains highly conserved regions with three dispersed AUUUA pentamers in an AU-rich context. Sequence analysis of the RET 3’UTR from a number of patients revealed a statistically significant overrepresentation of a SNP (rs3026785). This SNP causes a T to C conversion that is located in a U-rich tract near the last AUUUA pentamer and is associated with a protective haplotype (110). Reporter and mRNA decay experiments showed that the T to C conversion led to increased mRNA and reporter activity. The RET rs3026785 SNP is an example where polymorphism in an AU-rich region can act as a protective allele against other predisposing alleles, neutralizing their effects and moving carriers away from the disease-liability threshold (110).

4.1.3. Thrombin-activable fibrinolysis inhibitor

Thrombin-activatable fibrinolytic inhibitor (TAFI) is an enzyme that down-regulates fibrinolysis and is activated by thrombin, which is generated by the coagulation system. TAFI has also been implicated in the inactivation of the inflammatory mediators bradykinin, the anaphylatoxins C3a and C5a, annexin II, and osteopontin, suggesting roles in inflammation, wound healing, and blood pressure (111). High levels of TAFI increase the risk for venous thrombosis and inhibitors of TAFI are considered for treating bleeding disorders such as hemophilia (112). The 3’UTR of TAFI mRNA contains several SNPs; one particular SNP, rs1049669, located at position +1344, is a G to A substitution in a region that is AU-rich (UUGUUAUCA[A/G]AUUAAUUUAAGUUUAUU). The +1344 G/A substitution has a general destabilizing effect that could lead to lower TAFI expression and might influence the occurrence of vascular disease (113).

4.1.4. Human glycoprotein

Plasma cell membrane glycoprotein-1 (PC-1) inhibits insulin receptor, tyrosine kinase activity, and subsequently, insulin cellular signaling. Abrupt over-expression of PC-1 contributes to human insulin resistance (114). A haplotype in the 3’UTR of PC-1 mRNA was implicated in increased PC-1 mRNA stability and expression, which leads to increased risk for insulin resistance. This haplotype called “P” consists of a cluster of three closely spaced, single-nucleotide polymorphisms: G2897A, G2906C, and C2948T. The three nucleotide changes are localized between two AUUUA motifs. Although these elements do not represent typical AREs and the mutations do not directly disrupt them, the P haplotype leads to a two-fold increase in mRNA stability in reporter assays. A study showed that individuals carrying the “P” haplotype had higher PC-1 protein content in both skeletal muscle and cultured skin fibroblasts. They were also at higher risk for insulin resistance, had higher levels of plasma glucose and insulin during an oral glucose tolerance test, and had higher levels of cholesterol, HDL cholesterol, and systolic blood pressure. In another independent study, subjects with type 2 diabetes, the most likely clinical outcome of insulin resistance, had a higher P haplotype frequency than did healthy control subjects (115).

4.1.5. Human glucocorticoid receptor

Glucocorticoids are steroid hormones that are produced by the adrenal gland. They are secreted in stress situations and have a wide range of effects on metabolism, the immune system, and behavior. Glucocorticoids are widely used as anti-inflammatory drugs to treat diseases such as rheumatoid arthritis and asthma. They bind to an intracellular receptor called the glucocorticoid receptor (GR) (116). The human GR gene is alternatively spliced into two isoforms, hGR Alpha and hGR Beta. hGR Alpha and hGR Beta share the N-terminal 727 amino acids and hGR Alpha has 50 additional amino acids while the shorter hGR Beta has 15 distinct amino acids. Alternative splicing also leads to a reduction in the size of the 3’UTR of hGR Beta (about 1 kb less than hGR Alpha) (117-118). The hGR Alpha protein localizes to the cytoplasm, can bind to the glucocorticoid receptor, translocates to the nucleus, binds to the glucocorticoid response elements, and alters transcription of genes. Alternatively, it can regulate the activity of other transcription factors such as AP-1 or NF-kB. In contrast, hGRBeta does not bind to the receptor and is transcriptionally inactive; it can even act as a dominant negative inhibitor of hGR Alpha, contributing to glucocorticoid resistance in patients (119). Four AUUUA motifs are present in the 3’UTR of hGR Beta mRNA, and ten are located in the 3’UTR of hGR Alpha mRNA. A polymorphism (3669A>G) in hGR Beta 3’UTR is associated with rheumatoid arthritis; this A to G mutation in an AUUA motif (to GUUUA) results in an increased stability of mRNA and hGR Beta protein expression (118, 120).

4.2. Deletions/insertions AU-rich element function and disease

4.2.1. PPP1R3; a regulator of muscle glycogen

PPP1R3 gene encodes RGI1, a 124-kDa regulatory subunit of skeletal muscle, glycogen-associated, type 1 serine-threonine protein phosphatase (PP1G). RGI1 is expressed in heart and skeletal muscles and regulates the activity of PP1G affecting the balance between glycogen synthesis and glycogenolysis (121). Defective insulin-mediated activation of muscle glycogen synthase activity is a risk factor for type 2 diabetes, a
metabolic disorder that is characterized by high blood glucose in the context of insulin resistance and relative insulin deficiency (122). Insulin resistant Pima Indians of Arizona who have the world’s highest reported prevalence of type 2 diabetes are characterized by decreased activity of PP1G (123). An investigation of the PPP1R3 gene in Pima Indians led to the discovery of two insertion-deletion polymorphism common variants in an AU-rich element in the 3’UTR. (124). The 5-bp deletion leads to the reduction of the distance between two AUUUA pentamers from ten (ARE1 allele) to only two base pairs (ARE2 allele) leading to lower mRNA stability, and subsequently, to decreased levels of RG1 protein (124). These events were associated with insulin resistance and type 2 diabetes in the Pimas. The same deletion polymorphism was investigated in the Canadian Oji-Cree population. The 3’UTR variation was significantly associated with variation in two-hour postprandial glucose results in adult Oji-Cree with type 2 diabetes or impaired glucose tolerance. Oji-Cree with diabetes or impaired glucose tolerance who were ARE2/ARE2 homozygotes had significantly lower two-hour postprandial plasma glucose than subjects with the other genotypes (125). The frequency of ARE1 and ARE2 was also determined in a study of 426 Japanese type 2 diabetic and 380 nondiabetic subjects. The allele frequency of the ARE2 variant in diabetic subjects was higher than that in nondiabetic subjects (0.34 vs. 0.29), even though its frequency in Japanese subjects was lower than the reported value in Pima Indians (0.56) (126). The PPP1R3 3’UTR polymorphism is also associated with clinical and hormonal characteristics of women with polycystic ovary syndrome, which can be caused by insulin resistance (127). The 5-bp deletion/insertion polymorphism in the ARE of PPP1R3 is the most investigated of its kind, affecting ARE function with four independent reports analyzing its relevance in type 2 diabetes in three different ethnic groups or on women with polycystic ovary syndrome (124-127).

4.2.2. Beta 2-adrenergic receptor

Beta 2-Adrenergic receptors (Beta 2-AR) are G protein-coupled receptors that are expressed on airway epithelial and smooth muscle cells. The Beta 2-ARs dilate smooth muscle cells, thereby regulating mucociliary clearance and relaxation. Beta 2-AR agonists are the most effective bronchodilators available, and give rapid relief of asthma symptoms (128-129). Three haplotypes of Beta 2-ARs were found that differ in the number of C repeats in a poly-C tract in the 3’UTR: 11, 12, or 13 C repeats. The overall sequence of Beta 2-AR 3’UTR is rich in A and U and the poly-C region lies only 20 bp downstream of an AUUUA pentamer. Actinomycin-D chase experiments revealed that a reporter with the 11C haplotype had a significantly lower mRNA half-life compared to the other two haplotypes. The polymorphism does not directly affect an ARE sequence and the mRNA destabilizing effect might be due to the reduction of the distance between two AU-rich regions. This heterogeneity may contribute to the variability of clinical responses to Beta-agonist, and genotyping to identify these 3’UTR polymorphisms may improve predictive power within the context of Beta 2-AR haplotypes in pharmacogenetic studies (130).

4.2.3. Low-density lipoprotein Receptor

Low-Density Lipoprotein Receptor (LDLR) is a cell surface receptor that mediates the endocytosis of cholesterol-rich LDL. LDLR-associated familial hypercholesterolemia is the most frequent Mendelian disorder and is a major risk factor for the development of coronary artery diseases (131-132). Familial hypercholesterolemia Helsinki is a deletion in the LDLR gene that deletes 9.6 kb from intron 15 to exon 18 (133). The transcript from the gene with the deletion can be alternatively spliced into two forms that lack the membrane-spanning domain and the cytoplasmic domain, and thus, are defective in the internalization of LDL. This results in accumulation of LDL in the circulation and produces the clinical picture of familial hypercholesterolemia. The amounts of each of the two mutant transcripts were approximately 10 times higher than that of the normal transcript in a heterozygous patient. The normal LDLR 3’UTR contains two potent AREs: UUUUAUAUUUAAU and UUAUUAAUUUUA, which are targets of RNA-binding proteins such as hnRNP D, hnRNPI, and KSRP (134). Both regions are deleted in the familial hypercholesterolemia Helsinki-carrying allele, which might explain the higher abundance of the mutants due to higher mRNA stability. This is an example where a deletion that includes AU-rich elements leads to higher expression of a truncated and defective protein (135).

4.2.4. Thymidylate Synthase

Thymidylate synthase (TS) is essential for the biosynthesis of thymidine monophosphate (dTMP), which is subsequently phosphorylated to thymidine triphosphate for use in DNA synthesis and repair. TS is a major target for cancer chemotherapeutic drugs, especially the widely used 5-fluorouracil (136). The screening of expressed sequence tags (EST) lead to the identification of a candidate 6-bp (UUAAAG) deletion (D)/insertion (I) variation at bp 1494 in the 3’UTR TS mRNA. The allele frequency of the 6-bp deletion was found to be 0.29 in a Caucasian population (137). This polymorphism was found to be associated with inflammatory disease, cancer, and developmental malformations (138-140). Reporter assays show that TS mRNA is stable and contains no detectable elements that cause significant transcript instability or translational silencing. However, the 6 bp deletion leads to increased decay of a reporter message and could be linked to low TS mRNA levels in vivo (141). Interestingly, the mRNA decay-promoting protein AUF1 displayed preferential affinity for the D allele leading to its rapid decay (139). The D allele is an atypical binding partner for AUF1; it is relatively AU-rich but does not contain typical ARE pentamers. The differential regulation of the D allele by AUF1 is one of few examples that link the binding of an RNA binding protein (RBP) to a mutation in a non-coding RNA, underscoring the critical roles of RBPs as post-transcriptional genetic determinants in health and disease.
Sequence variations in AU-rich elements and disease

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Sequence variation affecting ARE</th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUF1</td>
<td>ARE-binding, mRNA destabilization</td>
<td>Alternative Splicing</td>
<td>N.I.</td>
</tr>
<tr>
<td>BMP2</td>
<td>Osteogenesis</td>
<td>SNP</td>
<td>Osteoporosis, poor fitness</td>
</tr>
<tr>
<td>CCND1</td>
<td>Cyclin D1; cell cycle regulation</td>
<td>Alternative Splicing, Alternative Polyadenylation</td>
<td>Mantle Cell Lymphoma</td>
</tr>
<tr>
<td>CTKA</td>
<td>Casein kinases, signaling, diverse cellular processes</td>
<td>Alternative Polyadenylation</td>
<td>N.I.</td>
</tr>
<tr>
<td>COX-2</td>
<td>Prostaglandin Synthase, inflammation</td>
<td>Alternative Polyadenylation</td>
<td>Colorectal adenocarcinoma</td>
</tr>
<tr>
<td>FGF-2</td>
<td>Growth factor; embryogenesis, wound healing</td>
<td>Alternative Polyadenylation</td>
<td>ARE lacking form predominant in transformed cells</td>
</tr>
<tr>
<td>FLT1</td>
<td>VEGF receptor</td>
<td>Alternative Polyadenylation</td>
<td>N.I.</td>
</tr>
<tr>
<td>G0R</td>
<td>Glucocorticoid Receptor</td>
<td>SNP</td>
<td>Rheumatoid Arthritis</td>
</tr>
<tr>
<td>HnR</td>
<td>ARE-binding, mRNA stabilization</td>
<td>Alternative Polyadenylation</td>
<td>N.I.</td>
</tr>
<tr>
<td>LDLR</td>
<td>Low-Density Lipoprotein Receptor, endocytosis of cholesterol-rich LDL</td>
<td>Deletion</td>
<td>Familial Hypercholesterolemia</td>
</tr>
<tr>
<td>PC-1</td>
<td>Plasma cell membrane glycoprotein-1; inhibits insulin cellular signaling</td>
<td>SNP</td>
<td>Diabetes, high cholesterol, high systolic blood pressure</td>
</tr>
<tr>
<td>PHLFCPS1</td>
<td>Phosphducin-like protein; G-protein signaling</td>
<td>Alternative Polyadenylation</td>
<td>N.I.</td>
</tr>
<tr>
<td>PPP1R3</td>
<td>Codes for RG1; affects balance between glycolgen synthesis and glycojenolysis</td>
<td>5bp deletion</td>
<td>Type 2 Diabetes</td>
</tr>
<tr>
<td>RET</td>
<td>Receptor for the glial-cell-line-derived neurotrophic factor</td>
<td>SNP</td>
<td>Protective allele against Hirschprung disease</td>
</tr>
<tr>
<td>STAR</td>
<td>Stereoidogenic Acute Regulatory protein, elevation of cholesterol metabolism</td>
<td>Alternative Polyadenylation</td>
<td>N.I.</td>
</tr>
<tr>
<td>TAF1</td>
<td>Thrombin Activatable Fibrinolysis Coagulation</td>
<td>SNP</td>
<td>Possible implications in vascular disease</td>
</tr>
<tr>
<td>TCR Zeta</td>
<td>T cell Receptor amplification module</td>
<td>Alternative Splicing</td>
<td>Systemic Lupus Erythematosus</td>
</tr>
<tr>
<td>TS</td>
<td>Thymidylate synthase; biosynthesis of thymidine monophosphate</td>
<td>6bp deletion</td>
<td>Inflammatory Lupus Erythematosus, cancer and developmental malformations</td>
</tr>
<tr>
<td>Beta 2-AR</td>
<td>Beta2-Adrenergic receptors G; coupled receptor expressed on airway</td>
<td>Variable-length poly-C tract polymorphism</td>
<td>May contribute to the variability of clinical responses to Beta-agonist</td>
</tr>
<tr>
<td>Beta- catenin</td>
<td>Wnt signaling</td>
<td>Alternative Splicing</td>
<td>Esophageal Cancer</td>
</tr>
</tbody>
</table>

N.I. Not Investigated

5. PERSPECTIVES

Many AREs are present in genes that perform highly regulated functions such as cytokines and cell cycle components. Such elements must be highly conserved and it is unlikely that they carry genetic mutations (9, 142). In fact, genetic variations in type II ARE sequences that contain three or more clustered and overlapping AUUUA pentamers may not exist at all. Loss-of-function mutations would lead to abnormally elevated expression of corresponding genes and are expected to be detrimental. For instance, the deletion of the ARE from the TNF-Alpha gene in mice leads to severe pathological disorders like chronic inflammatory arthritis and Crohn's-like inflammatory bowel disease (143). Most of the discovered sequence variations affect AREs of scattered pentamers or ARE-like regions and do not always lead to stabilization of the affected mRNA. A comprehensive literature search revealed twenty genes with ARE-variations that are relevant to disease (Table 1). Interestingly, one of the most investigated genetic variations with proven clinical relevance in insulin resistance is a gain-of-function deletion in the PPP1R3 gene that brings two AUUUA pentamers to closer proximity leading to a reduction in mRNA stability and protein expression (124-127).

Mutations and/or aberrant alternative splicing or polyadenylation in or near ARE regions in somatic cells are expected to be more common than genetic polymorphisms. Such somatic variations are reported in a number of genes like cyclin D1, TCR zeta chain, FGF-2, and COX-2 (29, 51, 67). In fact, a recent finding suggests that shortening of 3’UTRs by alternative cleavage and polyadenylation is wide spread in cancer cells (144); therefore, it is likely to find novel ARE-containing genes that are aberrantly regulated due to this phenomenon.

Surprisingly, some reports indicate that mRNA variants with AU-rich motifs lead to the stabilization rather than the destabilization of the corresponding mRNA. In the case of cyclin D1 mRNA in Mantel cell lymphoma, the destabilization is suggested to be the result of the activity of miRNAs such as the miR-15/16 family and the miR17-92 cluster, while the AU-rich motifs lead to stabilization (31). In addition, the AU-rich mRNA variant of the T-cell receptor zeta chain in systemic lupus erythematosus has a higher stability that the shorter variants (52). ARE-containing mRNAs are expected to have higher stabilities in inflammation or in tumor cells, probably because of aberrant signaling pathways that target AREs but fail to influence miRNA activity (2, 145).

This review has focused on functional and disease consequences of sequence variations in, or in close proximity to, AREs. It should be noted, however, that cross-talk between AREs and other non-AUUUA regulatory elements such as miRNA targets sites in the regulation of stability, and translation of a number of mRNAs has been reported (146-147). Therefore, it is likely that sequence variations in other cis acting regulatory elements, especially miRNA sites, might influence ARE function.
Sequence variations in AU-rich elements and disease

6. REFERENCES


Sequence variations in AU-rich elements and disease


34. C. Zhou, C. Z. Vignere and E. S. Levitan: AUF1 is upregulated by angiotensin II to destabilize cardiac Kv4.3 channel mRNA. *J Mol Cell Cardiol*, 45(6), 832-8 (2008)


1855


Sequence variations in AU-rich elements and disease


75. H. Duan and C. R. Jefcoate: The predominant cAMP-stimulated 3 x 5 kb StAR mRNA contains specific sequence elements in the extended 3'UTR that confer high basal instability. *J Mol Endocrinol*, 38(1-2), 159-79 (2007)


Sequence variations in AU-rich elements and disease


99. T. J. Yong, Y. Y. Gan, B. H. Toh and J. W. Sentry: Human CKI alpha(L) and CKI alpha(S) are encoded by both 2.4- and 4.2-kb transcripts, the longer containing multiple RNA-destabilising elements. Biochim Biophys Acta, 1492(2-3), 425-33 (2000)


Sequence variations in AU-rich elements and disease


Send correspondence to: Khalid S. A. Khabar, King Faisal Specialist Hospital and Research Center, P3354, MBC-03, Takhasusi Road, Riyadh 11211, Saudi Arabia, Tel: 966-1-442-7876, Fax: 966-1-442-4182, E-mail: khabar@kfshrc.edu.sa

http://www.bioscience.org/current/vol17.htm