January 2011

Apobec-1 Complementation Factor (ACF) Binds and Regulates Multiple RNAs

Kimberly Delaney
Washington University in St. Louis

WASHINGTON UNIERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences

Molecular Cell Biology

Dissertation Examination Committee:
Kathleen B. Hall, Chair
Nicholas O. Davidson
John Majors
Aubrey Morrison
Deborah Rubin
Sheila Stewart
Jason Weber

APOBEC-1 COMPLEMENTATION FACTOR (ACF) BINDS AND REGULATES MULTIPLE RNAs

By

Kimberly J. Delaney

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2011

Saint Louis, Missouri

ABSTRACT OF DISSERTATION

APOBEC-1 COMPLEMENTATION FACTOR (ACF) BINDS AND REGULATES MULTIPLE RNAs

By

Kimberly Jo Delaney

Doctor of Philosophy in Biomedical Sciences (Molecular Cell Biology)

Washington University in St. Louis, May 2011

Senior Professor Kathleen B. Hall, Advisor

AU-Rich Elements (ARE) are *cis*-acting RNA sequences in the 3'UTRs of a wide range of transcripts that function to regulate mRNA stability, localization, and/or translation through interaction with ARE-Binding Proteins. Apobec-1 Complementation Factor (ACF) was originally identified as a co-factor in ApoB mRNA C to U editing, but has recently been implicated in regulation of mRNA stability as an ARE-Binding Protein. Here we have used tissue culture models with RNA turnover assays to show that the stability of reporters cloned upstream of the Interleukin-6 (IL-6) 3'UTR or portions of the Cox-2 3'UTR is regulated by levels of ACF expression. Surprisingly, ACF expression results in stabilization of a reporter associated with the IL-6 3'UTR while resulting in destabilization of a reporter associated with the Cox-2 3'UTR. In order to more fully probe this dual ability of ACF, we examined its behavior as an RNA-binding protein. Purified recombinant truncations of ACF were used to probe the affinity and specificity of ACF binding to a panel of RNAs including ApoB mRNA (its canonical target) as well as IL-6 and Cox-2 RNAs, which we have shown are regulated *in cellulo* by ACF expression The first 380 amino acids of ACF, which contain three RNA recognition motifs (RRM), bind IL-6 and Cox-2 RNAs with higher affinity than ApoB mRNA. This protein is also capable of binding a C/U-rich RNA (GABA Intron), indicating that ACF has a broader target preference than simply AU-Rich RNAs. Furthermore, *in vitro* binding assays reveal that RRM1 of ACF binds IL-6, Cox-2, and GABA Intron RNA but not ApoB, while RRM3 does not detectably bind any of the RNAs probed. This indicates that RRM1 participates in RNA binding of some targets but not others and RRM3 is not

part of the RNA-binding domain. These observations were extended using tryptophan fluorescence to determine that Cox-2 RNA interacts with ACF RRM2, suggesting that RRMs 1 and 2 together bind at least some RNA targets. UV-crosslinking assays identified discrete ACF binding sites within both the IL-6 3'UTR and the Cox-2 3'UTR. On both RNAs, these sites are consistent with regions that confer ACF-dependent mRNA stabilization or destabilization in RNA turnover assays. UV-crosslinking assays also revealed a structural preference in the ACF:Cox-2 interaction. Finally, these observations were examined in the context of ACF structural predictions. While little experimentally determined structural data exists for ACF, homology modeling was used to predict possible secondary and tertiary three-dimensional structures that may account for the physiological and binding activities observed. We propose that ACF binds RNA targets by multiple mechanisms, using one or more of its RRMs, and that the resulting complex displays the RNA to facilitate RNA editing, stabilization, or destabilization. We suggest that the geometry of the complex also impacts ACF interactions with co-factors such as Apobec-1 in mRNA editing or other ARE-Binding proteins that together regulate the stability of common RNA targets.

ACKNOWLEDGEMENTS

These may be the most difficult few pages of my thesis I have written. My time in graduate school as been both challenging and inspiring, and shocking as it may be to those who know me, I find it difficult to express my gratitude in words.

First, I want to thank Dr. Kathleen Hall. From the first time I met her (complete with a Blue-State streak in her hair) she has demanded nothing but the best from me as a student and scientist and offered nothing but the best in support, encouragement, and most importantly, teaching. She is a true teacher in every sense of the word and she has inspired me and restored my love of this challenging and often exasperating field of study. In many ways, literally and figuratively speaking, I would not have completed my thesis work without her belief in my abilities as a scientist and the power of a fresh start.

I cannot thank my current mentor without also thanking my past mentor, Dr. Arthur Hunt. He seemed to know I would get my Ph.D. before I knew it was an option. From my first year at the University of Kentucky he guided me, taught me, and gave me opportunities few undergraduate researchers ever experience. I truly appreciate all he has done as a PI, a mentor, and a friend to encourage my growth as a scientist.

It is also important for me to thank the faculty members on my committee that have helped to shape my thesis and my career as a scientist. They have been supportive and patient as they've watched my project evolve and helped me deal with the unique challenges my graduate school career has presented.

I also would like to thank the Hall Lab for all their support and patience during my tenure with them. When I joined this group of outstanding scientists I had so much to learn, but everyone was patient and generous and accepted me immediately. When I thought the ACF paper would never happen, Sandra Williams told me to "build it and it will come," so she built me a wall to post my data – and the data came soon after. For that wisdom and encouragement, I am immensely grateful. The Hall Lab has become more than colleagues, they are friends. The same can be said of my former labmates in Gastroenterology. During some of the best and worst times of my life personally and scientifically, Valerie Blanc, Susan Kennedy, and Libby Newberry offered support, congratulations, and shoulders to lean on. Valerie Blanc let me into her world of ACF

and shared with me her vast knowledge of this project. Her guidance and friendship have helped me navigate through the science and the stress of graduate school.

My deepest gratitude goes out to my friends who have been with me since the first day of classes here at Washington University. Their support has been a wonderful help during this endeavor. Specifically, I want to thank Sarah Naylor. She has been my constant friend and companion during victories and defeats both professional and personal. She has listened to more complaints and worries than anyone should ever hear and helped me every step of the way from choosing the next cloning scheme to finding a new home or the perfect outfit to wear for a talk. I hope I will be able to call her my friend and colleague for years to come and that I may somehow repay her for the gift of her support and friendship.

Finally, I need to thank my parents Ed and Sharon Delaney and my brother Chris Delaney. They have not always understood my desire to stay in school for nearly 23 years or the need to work nights and weekends to transfect cells or isolate RNA. However, they have lived and died with the success of every experiment. They have boosted me up and caught me when I've fallen, which has been hard and often. Their love and support is unmatched and will never be unappreciated. Few people are as blessed as I have been to have such amazing family to make the difficult process of graduate school a little easier. I know it could never have been easy for them to watch this seemingly endless process, and I hope they will accept this thesis as the smallest repayment for their incredible generosity and love.

This work has been supported by:

Division of Biology and Biomedical Sciences Molecular and Cell Biology NIH Training Grant

AHA Predoctoral Fellowship 910115

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

Introduction

RNA Editing.

While control of transcription is a key step in the restriction of gene expression, co- and post-transcriptional modifications and regulation are equally important in influencing the expression of protein. As a general rule, eukaryotic mRNAs are co-transcriptionally spliced, polyadenylated, and capped. These steps allow for layers of regulation in that alternative splicing results in altered protein isoforms and failure to polyadenylate or cap an mRNA results in degradation and failure to translate. However, there are further mRNA modifications that influence gene expression.

RNA editing is an important mRNA modification that allows tightly regulated variation of gene expression. RNA editing is generally defined as any post-transcriptional process that modifies an RNA transcript sequence from its template DNA or RNA. Editing of ribosomal and transfer RNAs has been observed as well as editing of protein-coding mRNAs [1]. There are two well-characterized forms of mRNA editing to date: A to I and C to U. A to I editing is mediated by adenosine deaminases acting on double stranded RNA (ADARs) [1]. This form of editing has an absolute requirement for double stranded RNA (dsRNA) and only occurs on pre-mRNA templates containing intronic regions. ADARs contain both dsRNA binding domains and catalytic deaminase domains, allowing them to function independently without the assistance of co-factors [2, 3]. A to I editing occurs in all organisms from *C. elegans* to mammals; the primary mRNAs that undergo A to I editing include voltage- and ligand-gated ion channel membrane proteins in nervous tissue as well as ADAR transcripts themselves [3].

The other well-characterized form of editing involves deamination of cytidine to uridine (C to U). The most well-known form of C to U editing occurs in the spliced and

polyadenylated nuclear apolipoproteinB mRNA transcript, which is over 14,000 bases long. C to U editing of base 6666 results in an in-frame mutation of CAA, which codes for Gln, to UAA, a stop codon. This mutation results in the production of a truncated protein variant, referred to as ApoB 48, while unedited protein is referred to as ApoB 100 (Figure 1.1A). Both of these proteins participate in lipid trafficking, but have very different physiological roles. Editing has been observed in the small intestine of mammals (>85% of apoB mRNA is edited) as well as the liver of some small species, specifically rodents [4].

ApoB mRNA Editing.

Unlike A to I editing, ApoB mRNA is edited primarily after splicing and polyadenylation but before nuclear export [5], and mediated by ApoB mRNA editing catalytic subunit 1 (Apobec-1). This highly conserved protein functions as a dimer to both bind RNA and deaminate a specific cytidine residue, thus resulting in a premature stop codon. Mutagenesis has revealed that both the RNA binding and deamination functions of Apobec-1 are necessary for editing activity [6].

Extensive RNA mutagenesis and *in vitro* editing assays have determined that ApoB mRNA editing requires a minimal sequence of approximately 30 nucleotides that flanks the edited cytidine (Figure 1.1B) [7, 8]. This region includes an 11 nucleotide mooring sequence that is 4-6 nucleotides downstream of the edited base [7, 8], and an AU-rich Apobec-1 consensus binding site which is three nucleotides downstream of the edited base: UUUN(A/U)U [9]. There are also 5' and 3' efficiency elements outside the 30 nucleotide region which are necessary to achieve *in vivo* levels of editing [10, 11].

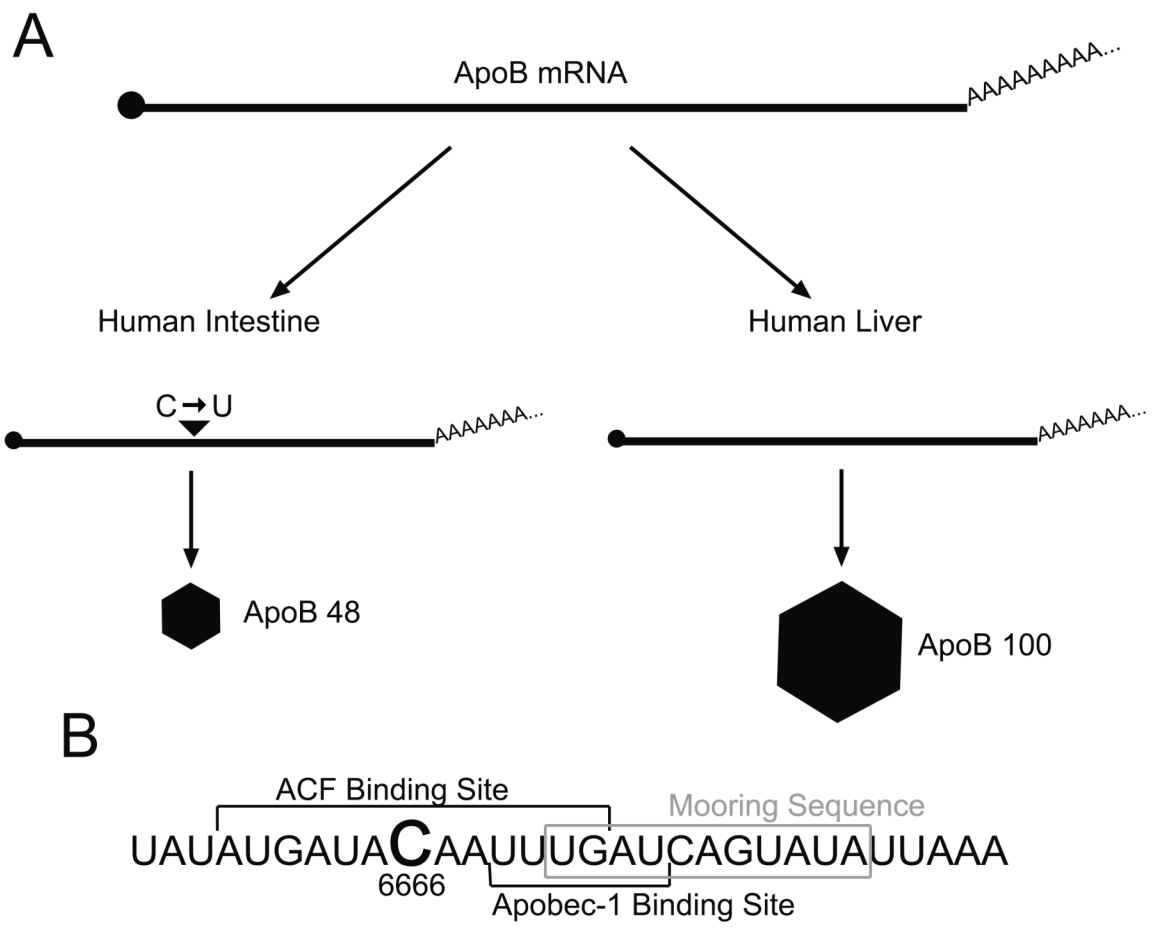


Figure 1.1 - ApoB mRNA is edited to produce two gene products. A) Schematic of ApoB mRNA editing to produce ApoB 48 and ApoB 100. B) 30 nt editing cassette of human ApoB mRNA with edited C6666 in bold.

RNA folding prediction algorithms and RNase mapping suggest that this region adopts a stem-loop secondary structure that is important for editing efficiency [12, 13].

Soon after the discovery of edited ApoB mRNA, it was estimated that the editing enzyme comprised a total molecular weight of 120-125 kDa [14, 15] and the editosome resides in a 27S complex when isolated by glycerol gradient sedimentation [16]. This is consistent with findings upon the cloning of Apobec-1 that this deaminase alone is not sufficient for RNA editing, but must be supplemented by some type of complementation factor(s) [17]. This complementation factor was identified as a 65 kDa RNA binding protein isolated using ApoB RNA as bait [18, 19]. This protein was later identified as Apobec-1 Complementation Factor (ACF), found on human chromosome 10 over 15 exons [20]. Together, ACF and Apobec-1 are sufficient for synthetic editing [21].

Apobec-1 Complementation Factor

While ApoB mRNA editing has only been observed in the human small intestine, ACF is expressed in a wide array of human tissues including high expression in the liver, colon, kidney, spleen, brain, and many others [22, 23]. Examination of ESTs has suggested up to fifteen different splice variants of ACF. However, only a few have been observed to function *in vivo*: ACF 65 (full length) and ACF 64 differ by an eight amino-acid deletion in exon 12, though both appear to fully complement editing in cells [23]. ACF 43 and ACF 45 result from alternative polyadenylation and splicing in exon 11, with both proteins encoding amino acids 1-379 of ACF with C-terminal additions of 4 and 26 amino acids, respectively [24]. These variants are tissue- and species-specific and

are able to bind ApoB RNA and complement Apobec-1 editing of ApoB mRNA, though with decreased efficiency.

ACF exhibits significant sequence homology to the hnRNP family of proteins, particularly in that it contains three predicted RNA Recognition Motif (RRM) domains between amino acids 58 and 303. RRMs 1 and 3 exhibit homology to the well studied RNA binding RRM proteins Sex Lethal and HuD (Figure 1.2). ACF mutagenesis, along with the activity of the ACF 43 and ACF 45 splice variants, indicates that ApoB RNA binding is isolated to the first 380 amino acids of ACF, though the activity can likely be isolated further with more extensive study [24, 25]. There is a putative double stranded RNA Binding Domain (dsRBD) that extends from amino acids 438 to 515, though it is not necessary to complement ApoB mRNA editing and to date no function has been observed [25]. Finally, ACF contains a novel nuclear localization signal (ANS), which confers nuclear-cytoplasmic shuttling, though editing is predominantly in the nucleus [21]. Phosphorylation is another potential regulator of ACF sub-cellular localization. Two serines (S154 in putative RRM2 and S368) are phsophorylated by protein phosphatase 1 (Figure 1.2B) [26]. However, Phospho-ACF is only detectable in the nucleus, while cytoplasmic ACF appears to be un-phosphorylated [4].

Extensive mutagenesis of both ACF and ApoB mRNA has revealed much about the ACF:ApoB interaction. The ACF binding site on the ApoB mRNA has been narrowed down to a 13 nucleotide AU-rich stretch flanking the edited base: AUGAUA**C**AAUUUG (Figure 1.1B) [27]. The domain of ACF that interacts with ApoB mRNA has not been specifically identified, though it has been narrowed down to amino acids 150-380, while the Apobec-1 interacting domain appears to extend from amino acid

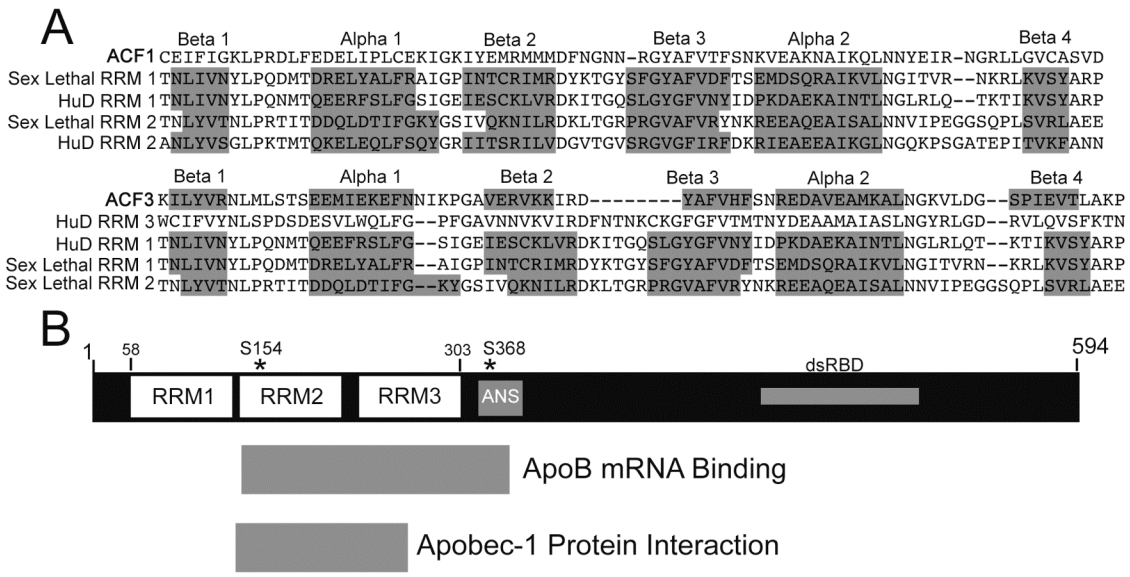


Figure 1.2 - ACF. A) Alignment of ACF RRMs 1 and 3 with the RRMs of homologous RNA Binding proteins: Sex Lethal and HuD. (ClustalW2, [28]). B) Schematic of ACF with putative domains labeled. RRM is RNA Recognition Motif. ANS is ACF Nuclear Signal. dsRBD is double stranded RNA Binding Domain. * indicates locations of phosphorylation. ApoB mRNA Binding and Apobec-1 protein interaction regions as defined by Blanc et al. (2001, [25]).

144 to 257 (Figure 1.2B) [25]. This would suggest that putative RRMs 2 and 3 may interact with RNA and/or Apobec-1.

Due to ACF's expression in a wide range of tissues beyond the scope of ApoB mRNA editing, a germline knockout of *Acf* was generated. However, these mice are embryonic lethal at day E3.5 [29]. While heterozygous mice are grossly normal with no observable phenotype at steady state, *Acf* null mice develop to the blastocyst stage but fail to implant. Isolated $Acf^{-/-}$ blastocysts also fail to proliferate *in vitro* compared to wild-type litter mates [29]. This lethality occurs at a stage in development when there is no detectable ApoB mRNA or Apobec-1 and no small intestine in which editing might occur. These observations suggest that ACF plays a highly important role beyond its participation in ApoB mRNA editing. As reported by Blanc et al (2010), ACF has an alternative function as an AU-Rich Element binding protein (ARE-BP) that stabilizes interleukin-6 (IL-6) mRNA in Kupffer cells from mouse liver through interaction with the mRNA 3'UTR [30].

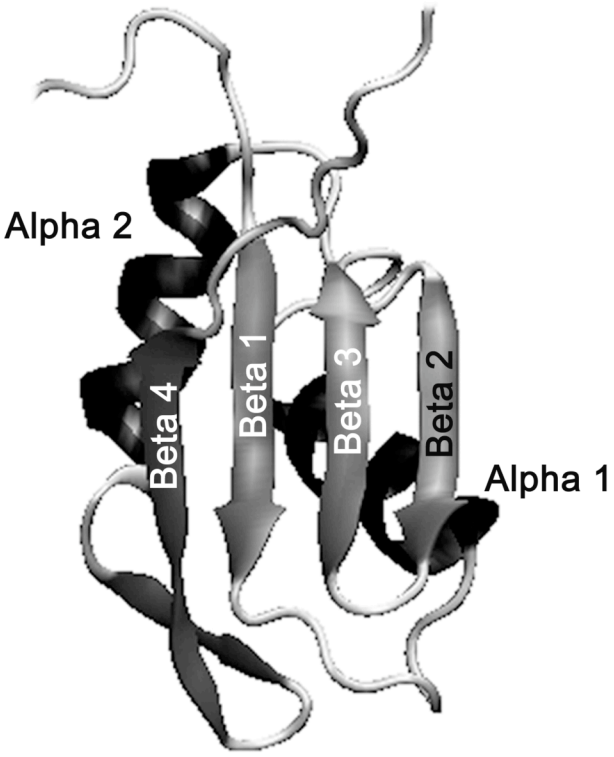
RNA Recognition Motifs (RRMs).

ACF contains three putative RRMs, based on sequence analysis. RRMs are present in 0.5-1% of human genes, making them one of the most abundant RNA binding domains in higher organisms [31, 32]. These domains and the proteins in which they are found have been widely studied structurally, biochemically, and functionally: RRMs are involved in nearly every co- and post-transcriptional process, including mRNA and rRNA processing, RNA export, and RNA metabolism [33, 34]. RRMs are approximately

70-90 amino acids, and consist of β-α-β-β-α-β topology. The four β-strands form an anti-parallel β-sheet supported from behind by two α-helices (Figure 1.3).

Canonically, 2-8 nucleotides of single-stranded RNA (ssRNA) are bound on the β-sheet [35-37]. While the RRM domain structure is generally conserved, the RNP motifs, found on β1 and β3, are the most highly conserved sequences within these domains (Figure 1.3) [38]. In both of these sequences, aromatic residues appear to play an important role in RNA contact, often by stacking with the RNA bases. However, these are not 100% conserved and are also not the only points of RRM:RNA contact. These aromatic residues can vary among RRMs and thus create higher degrees of sequence-specificity. For example, in the case of the RRM protein U1A, a Gln at RNP1 position 3 (canonically a Phe or Tyr) appears to be involved in sequence-specific hydrogen bonding with its target: U1 snRNA stem-loop II [39]. U1A is also a prime example of alternate points of contact in RRM:RNA interactions: in this RRM, loop 3 between β-strand 2 and β-strand 3 is necessary for RNA binding and substitution of this loop can result in abrogation of binding or decreased sequence-specificity [40-42].

While some RRM proteins exhibit extremely high target specificity, as in the case of U1A, which binds a single target with high affinity, many other RRMs have wider specificity. HuR is a well-studied RNA binding protein involved in mRNA metabolism and translation that contains three RRMs. Analysis of HuR's endogenous targets (via RNA immunoprecipitation) indicated that HuR binds primarily U-rich sequences with a loose 17-20 nucleotide consensus motif [43]. However, further analysis revealed that while HuR appears to exhibit high affinity for this consensus, a number of validated endogenous HuR targets don't contain this sequence. Clearly, while some RRM proteins



RNP 1: [RK] - [G] - **[FY]** - [GA] - **[FY]** - [ILV] - [X] - **[FY]**
Beta Strand 3

RNP 2: [ILV] - **[FY]** - [ILV] - [X] - [N] - [L]
Beta Strand 1

Figure 1.3 - The RNA Recognition Motif. Model of ACF 3 (pdb: 2cpd) [44]. Canonical RNP motifs listed with aromatic residues in bold.

demonstrate specific binding activity, other RRMs are capable of more generic interaction defined by the secondary structure or overall nucleotide composition of the RNA.

AU-Rich Element mRNAs.

Regulation of mRNA degradation is a powerful mechanism in the control of gene expression. It has been extensively observed that AU-rich elements (AREs) in the 3'UTRs of a wide range of mRNAs are cis-acting elements that work in concert with trans-acting RNA binding proteins to regulate mRNA stability and translation [45]. Originally identified in cytokines, AREs are estimated to be present in 10-15% of the transcriptome, including RNAs that regulate inflammatory and immune response, transcription, cellular proliferation, RNA metabolism, development, signaling, and tumor growth factors [46, 47]. The definition of an ARE ranges from the vaguely-defined "AU-rich or U-rich sequence" to the highly defined AUUUA pentamer or WWWUAUUUAUWWW (where W is U/A), a computationally defined consensus motif that was used as the building block for the ARE-mRNA database (ARED), which attempts to identify and classify AREs in the human genome [46, 48-50]. These RNA sequences appear to be targets for RNA binding proteins: binding can result in RNA stabilization, degradation, relocation, or translational control (reviewed in [47]). Most ARE mRNAs are bound in cell-, environment-, and stress-specific manners by multiple ARE-BPs [51-53]. Specifically, IL-6 and Cox-2 mRNAs are targets of Auf1, HuR, and TTP, as well as other ARE-BPs [54-61]. These two transcripts are highly regulated in the liver and colon, respectively, two regions of high ACF expression [22].

IL-6 is a cytokine and tumor growth factor that is associated with aging, cancer, liver regeneration, and the regulation of the inflammatory transcriptome [62-65]. The IL-6 mRNA has a highly AU-rich 3'UTR of 414 nt with five canonical ARE motifs (AUUUA), and is destabilized by AUF1 in an AU-rich dependent manner [57] but stabilized by cytoplasmic HuR [58]. Recently our work in Blanc et al (2010) showed that ACF modulates liver regeneration, possibly through post-transcriptional regulation of IL-6 mRNA. Specifically, we determined that ACF can directly bind a 129 nt segment of the IL-6 3'UTR and has the ability to stabilize endogenous IL-6 mRNA in murine Kupffer cells [30].

Cox-2 is the rate-limiting step in prostaglandin E2 synthesis, which is secreted from macrophages and plays an important role in T-cell activation, making Cox-2 vital in the regulation of aging, cardiac disease, and tumor growth, specifically in the breast and colon, a region of high ACF expression [22, 59, 66-68]. The 3'UTR of Cox-2 is 2.2 kB and contains 12 canonical ARE motifs. However, seven of these AREs are found in the first 64 nt of the 3'UTR, which is 85% AU-rich. Extensive work has shown that Cox-2 is regulated both transcriptionally and post-transcriptionally; it is stabilized by HuR interaction within the first 64 nt of the 3'UTR, while it is destabilized by both AUF1 and TTP [54-56, 60]. However, no interaction between ACF and Cox-2 RNA has previously been described.

ARE-Binding Proteins.

Though AREs are the *cis*-acting elements that confer reduced stability on the mRNA, this destabilization occurs through *trans*-acting factors: ARE-BPs. While there

are many ARE-BPs and more are continually being identified, a few have been extensively analyzed in their ability to bind RNA and modulate its stability. Of these, a number of protein families are represented: HuR, TIA-1, and Auf1 are RRM-containing proteins homologous to the hnRNP family of proteins; TTP is a zinc-finger protein; and KSRP is a KH-domain protein. All of these are RNA binding proteins that have been observed to regulate mRNA stability and/or translation.

HuR, Auf1, and TIA-1 are expressed ubiquitously and TTP is expressed very widely, though expression significantly varies between tissues. It is also interesting that expression of these proteins appears to be developmentally regulated as HuR and Auf1 expression increases with age while TTP and TIA-1 are expressed most highly in fetal tissue and drop with maturation [69]. Similarly, ACF is expressed widely, though varying among tissues [22], and certainly appears to function in a developmentally-regulated manner as suggested by the embryonic lethality phenotype. The diverse expression of these ARE-BPs both in tissue- and developmentally-specific manners contributes to the varied regulation of their targets in different tissues and stages of development.

What is particularly interesting about these ARE-BPs is that many share a common pool of RNA targets. For example, HuR stabilizes iNOS RNA which is destabilized by both TTP and KSRP. However, only HuR and KSRP interact with the RNA directly. While TTP is capable of binding RNA, it appears to regulate iNOS RNA *through* interaction with KSRP, which competes with HuR for RNA binding [70]. This is an example of a particularly intricate form of regulation in which expression of all three proteins results in both cooperation and competition to influence the stability of iNOS RNA.

Similarly, HuR and Auf1 share a large pool of target RNAs, some of which they can bind simultaneously but regulate differentially. In particular, both are capable of binding p21 and cyclin D1 RNAs concurrently, but HuR results in stabilization of both RNAs while Auf1 results in destabilization of both RNAs. Association of the proteins with these RNAs appears to be contingent on environmental stimuli, as UVC irradiation results in increased HuR:p21 complexes but decreased HuR:cyclin D1 complexes. However, UVC irradiation results in decreased Auf1:p21 complexes but increased Auf1:cyclin D1 complexes [71]. Alternatively, p16 mRNA is bound by both HuR and Auf1 at a common site on the mRNA's 3'UTR. These two proteins interact in an RNA-dependent manner and appear to bind this specific RNA cooperatively and together result in destabilization [72]. This is one of the few examples of HuR promoting mRNA decay; it is also quite the opposite of the combined effect of HuR and Auf1 on p21 and cyclin D1 mRNA stability in which the two proteins work in opposing manners. This clearly exemplifies the complexity of ARE-BP:RNA interactions *in vivo*. It is certain that the net effect of ARE-BP binding on mRNA stability is defined by the whole complement of proteins that bind cooperatively, concurrently, or competitively to the 3'UTR.

ACF's structural homology to many ARE-BPs, wide tissue-expression [22], clear role beyond ApoB mRNA editing [29], and preference for AU-rich RNA [27] all led us to predict that ACF can function as an ARE-BP. We hypothesize that ACF binds AU- or U-rich RNAs to regulate their stability and/or translation. Therefore, we began by expanding our findings from Blanc et al. (2010, [30]) and identifying ACF's behaviors in binding RNA and modulating the mRNA stability of both IL-6 RNA as well as a novel target: Cox-2 RNA.

REFERENCES

1 J. M. Gott and R. B. Emeson. (2000) Functions and mechanisms of RNA editing, Annu Rev Genet. **34**, 499-531

2 A. P. Gerber and W. Keller. (2001) RNA editing by base deamination: more enzymes, more targets, new mysteries, Trends Biochem Sci. **26**, 376-384

3 P. H. Seeburg. (2002) A-to-I editing: new and old sites, functions and speculations, Neuron. **35**, 17-20

4 D. M. Lehmann, C. A. Galloway, M. P. Sowden and H. C. Smith. (2006) Metabolic regulation of apoB mRNA editing is associated with phosphorylation of APOBEC-1 complementation factor, Nucleic Acids Res. **34**, 3299-3308

5 P. P. Lau, W. J. Xiong, H. J. Zhu, S. H. Chen and L. Chan. (1991) Apolipoprotein B mRNA editing is an intranuclear event that occurs posttranscriptionally coincident with splicing and polyadenylation, J Biol Chem. **266**, 20550-20554

6 N. Navaratnam, T. Fujino, et al. (1998) Escherichia coli cytidine deaminase provides a molecular model for ApoB RNA editing and a mechanism for RNA substrate recognition, J Mol Biol. **275**, 695-714

7 R. R. Shah, T. J. Knott, J. E. Legros, N. Navaratnam, J. C. Greeve and J. Scott. (1991) Sequence requirements for the editing of apolipoprotein B mRNA, J Biol Chem. **266**, 16301-16304

8 J. W. Backus and H. C. Smith. (1991) Apolipoprotein B mRNA sequences 3' of the editing site are necessary and sufficient for editing and editosome assembly, Nucleic Acids Res. **19**, 6781-6786

9 S. Anant and N. O. Davidson. (2000) An AU-rich sequence element (UUUN[A/U]U) downstream of the edited C in apolipoprotein B mRNA is a high-affinity binding site for Apobec-1: binding of Apobec-1 to this motif in the 3' untranslated region of c-myc increases mRNA stability, Mol Cell Biol. **20**, 1982-1992

10 M. Hersberger and T. L. Innerarity. (1998) Two efficiency elements flanking the editing site of cytidine 6666 in the apolipoprotein B mRNA support mooring-dependent editing, J Biol Chem. **273**, 9435-9442

11 M. Nakamuta, A. Tsai, L. Chan, N. O. Davidson and B. B. Teng. (1999) Sequence elements required for apolipoprotein B mRNA editing enhancement activity from chicken enterocytes, Biochem Biophys Res Commun. **254**, 744-750

12 C. Maris, J. Masse, A. Chester, N. Navaratnam and F. H. Allain. (2005) NMR structure of the apoB mRNA stem-loop and its interaction with the C to U editing APOBEC1 complementary factor, RNA. **11**, 173-186

13 N. Richardson, N. Navaratnam and J. Scott. (1998) Secondary structure for the apolipoprotein B mRNA editing site. Au-binding proteins interact with a stem loop, J Biol Chem. **273**, 31707-31717

14 D. M. Driscoll and E. Casanova. (1990) Characterization of the apolipoprotein B mRNA editing activity in enterocyte extracts, J Biol Chem. **265**, 21401-21403

15 N. Navaratnam, R. Shah, D. Patel, V. Fay and J. Scott. (1993) Apolipoprotein B mRNA editing is associated with UV crosslinking of proteins to the editing site, Proc Natl Acad Sci U S A. **90**, 222-226

16 H. C. Smith, S. R. Kuo, J. W. Backus, S. G. Harris, C. E. Sparks and J. D. Sparks. (1991) In vitro apolipoprotein B mRNA editing: identification of a 27S editing complex, Proc Natl Acad Sci U S A. **88**, 1489-1493
17 B. Teng, C. F. Burant and N. O. Davidson. (1993) Molecular cloning of an apolipoprotein B messenger RNA editing protein, Science. **260**, 1816-1819
18 A. Mehta, S. Banerjee and D. M. Driscoll. (1996) Apobec-1 interacts with a 65-kDa complementing protein to edit apolipoprotein-B mRNA in vitro, J Biol Chem. **271**, 28294-28299
19 A. Mehta and D. M. Driscoll. (1998) A sequence-specific RNA-binding protein complements apobec-1 To edit apolipoprotein B mRNA, Mol Cell Biol. **18**, 4426-4432
20 J. O. Henderson, V. Blanc and N. O. Davidson. (2001) Isolation, characterization and developmental regulation of the human apobec-1 complementation factor (ACF) gene, Biochim Biophys Acta. **1522**, 22-30
21 V. Blanc, S. Kennedy and N. O. Davidson. (2003) A novel nuclear localization signal in the auxiliary domain of apobec-1 complementation factor regulates nucleocytoplasmic import and shuttling, J Biol Chem. **278**, 41198-41204
22 A. Mehta, M. T. Kinter, N. E. Sherman and D. M. Driscoll. (2000) Molecular cloning of apobec-1 complementation factor, a novel RNA-binding protein involved in the editing of apolipoprotein B mRNA, Mol Cell Biol. **20**, 1846-1854
23 H. Lellek, R. Kirsten, I. Diehl, F. Apostel, F. Buck and J. Greeve. (2000) Purification and molecular cloning of a novel essential component of the apolipoprotein B mRNA editing enzyme-complex, J Biol Chem. **275**, 19848-19856
24 M. P. Sowden, D. M. Lehmann, X. Lin, C. O. Smith and H. C. Smith. (2004) Identification of novel alternative splice variants of APOBEC-1 complementation factor with different capacities to support apolipoprotein B mRNA editing, J Biol Chem. **279**, 197-206
25 V. Blanc, J. O. Henderson, S. Kennedy and N. O. Davidson. (2001) Mutagenesis of apobec-1 complementation factor reveals distinct domains that modulate RNA binding, protein-protein interaction with apobec-1, and complementation of C to U RNA-editing activity, J Biol Chem. **276**, 46386-46393
26 D. M. Lehmann, C. A. Galloway, C. MacElrevey, M. P. Sowden, J. E. Wedekind and H. C. Smith. (2007) Functional characterization of APOBEC-1 complementation factor phosphorylation sites, Biochim Biophys Acta. **1773**, 408-418
27 V. Blanc, N. Navaratnam, J. O. Henderson, S. Anant, S. Kennedy, A. Jarmuz, J. Scott and N. O. Davidson. (2001) Identification of GRY-RBP as an apolipoprotein B RNA-binding protein that interacts with both apobec-1 and apobec-1 complementation factor to modulate C to U editing, J Biol Chem. **276**, 10272-10283
28 R. Chenna, H. Sugawara, T. Koike, R. Lopez, T. J. Gibson, D. G. Higgins and J. D. Thompson. (2003) Multiple sequence alignment with the Clustal series of programs, Nucleic Acids Res. **31**, 3497-3500

29 V. Blanc, J. O. Henderson, E. P. Newberry, S. Kennedy, J. Luo and N. O. Davidson. (2005) Targeted deletion of the murine apobec-1 complementation factor (acf) gene results in embryonic lethality, Mol Cell Biol. **25**, 7260-7269
30 V. Blanc, K. J. Sessa, S. Kennedy, J. Luo and N. O. Davidson. (2010) Apobec-1 complementation factor modulates liver regeneration by post-transcriptional regulation of interleukin-6 mRNA stability, J Biol Chem. **285**, 19184-19192
31 J. C. Venter, M. D. Adams, et al. (2001) The sequence of the human genome, Science. **291**, 1304-1351
32 A. Clery, M. Blatter and F. H. Allain. (2008) RNA recognition motifs: boring? Not quite, Curr Opin Struct Biol. **18**, 290-298
33 G. Dreyfuss, V. N. Kim and N. Kataoka. (2002) Messenger-RNA-binding proteins and the messages they carry, Nat Rev Mol Cell Biol. **3**, 195-205
34 C. Maris, C. Dominguez and F. H. Allain. (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression, FEBS J. **272**, 2118-2131
35 F. H. Allain, P. Bouvet, T. Dieckmann and J. Feigon. (2000) Molecular basis of sequence-specific recognition of pre-ribosomal RNA by nucleolin, EMBO J. **19**, 6870-6881
36 C. Johansson, L. D. Finger, L. Trantirek, T. D. Mueller, S. Kim, I. A. Laird-Offringa and J. Feigon. (2004) Solution structure of the complex formed by the two N-terminal RNA-binding domains of nucleolin and a pre-rRNA target, J Mol Biol. **337**, 799-816
37 S. R. Price, P. R. Evans and K. Nagai. (1998) Crystal structure of the spliceosomal U2B"-U2A' protein complex bound to a fragment of U2 small nuclear RNA, Nature. **394**, 645-650
38 R. J. Bandziulis, M. S. Swanson and G. Dreyfuss. (1989) RNA-binding proteins as developmental regulators, Genes Dev. **3**, 431-437
39 T. H. Jessen, C. Oubridge, C. H. Teo, C. Pritchard and K. Nagai. (1991) Identification of molecular contacts between the U1 A small nuclear ribonucleoprotein and U1 RNA, EMBO J. **10**, 3447-3456
40 P. S. Katsamba, M. Bayramyan, I. S. Haworth, D. G. Myszka and I. A. Laird-Offringa. (2002) Complex role of the beta 2-beta 3 loop in the interaction of U1A with U1 hairpin II RNA, J Biol Chem. **277**, 33267-33274
41 C. Oubridge, N. Ito, P. R. Evans, C. H. Teo and K. Nagai. (1994) Crystal structure at 1.92 A resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin, Nature. **372**, 432-438
42 R. C. Bentley and J. D. Keene. (1991) Recognition of U1 and U2 small nuclear RNAs can be altered by a 5-amino-acid segment in the U2 small nuclear ribonucleoprotein particle (snRNP) B" protein and through interactions with U2 snRNP-A' protein, Mol Cell Biol. **11**, 1829-1839
43 I. Lopez de Silanes, M. Zhan, A. Lal, X. Yang and M. Gorospe. (2004) Identification of a target RNA motif for RNA-binding protein HuR, Proc Natl Acad Sci U S A. **101**, 2987-2992
44 T. Nagata, Y. Muto, M. Inoue, T. Kigawa, T. Terada, M. Shirouzu and S. Yokoyama. (2005) Solution structure of the RNA recognition motif of human APOBEC-1 complementation factor, ACF, Protein Database, pdb.org.

45 C. Y. Chen and A. B. Shyu. (1995) AU-rich elements: characterization and importance in mRNA degradation, Trends Biochem Sci. **20**, 465-470
46 A. S. Halees, R. El-Badrawi and K. S. Khabar. (2008) ARED Organism: expansion of ARED reveals AU-rich element cluster variations between human and mouse, Nucleic Acids Res. **36**, D137-140
47 K. S. Khabar. (2010) Post-transcriptional control during chronic inflammation and cancer: a focus on AU-rich elements, Cell Mol Life Sci. **67**, 2937-2955
48 T. Bakheet, M. Frevel, B. R. Williams, W. Greer and K. S. Khabar. (2001) ARED: human AU-rich element-containing mRNA database reveals an unexpectedly diverse functional repertoire of encoded proteins, Nucleic Acids Res. **29**, 246-254
49 T. Bakheet, B. R. Williams and K. S. Khabar. (2003) ARED 2.0: an update of AU-rich element mRNA database, Nucleic Acids Res. **31**, 421-423
50 T. Bakheet, B. R. Williams and K. S. Khabar. (2006) ARED 3.0: the large and diverse AU-rich transcriptome, Nucleic Acids Res. **34**, D111-114
51 N. S. Levy, S. Chung, H. Furneaux and A. P. Levy. (1998) Hypoxic stabilization of vascular endothelial growth factor mRNA by the RNA-binding protein HuR, J Biol Chem. **273**, 6417-6423
52 N. Xu, C. Y. Chen and A. B. Shyu. (2001) Versatile role for hnRNP D isoforms in the differential regulation of cytoplasmic mRNA turnover, Mol Cell Biol. **21**, 6960-6971
53 W. Wang, J. L. Martindale, X. Yang, F. J. Chrest and M. Gorospe. (2005) Increased stability of the p16 mRNA with replicative senescence, EMBO Rep. **6**, 158-164
54 S. J. Cok, S. J. Acton and A. R. Morrison. (2003) The proximal region of the 3'-untranslated region of cyclooxygenase-2 is recognized by a multimeric protein complex containing HuR, TIA-1, TIAR, and the heterogeneous nuclear ribonucleoprotein U, J Biol Chem. **278**, 36157-36162
55 S. J. Cok, 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**, 8196-8205
56 S. Sengupta, B. C. Jang, M. T. Wu, J. H. Paik, H. Furneaux and T. Hla. (2003) The RNA-binding protein HuR regulates the expression of cyclooxygenase-2, J Biol Chem. **278**, 25227-25233
57 S. Paschoud, A. M. Dogar, C. Kuntz, B. Grisoni-Neupert, L. Richman and L. C. Kuhn. (2006) Destabilization of interleukin-6 mRNA requires a putative RNA stem-loop structure, an AU-rich element, and the RNA-binding protein AUF1, Mol Cell Biol. **26**, 8228-8241
58 H. Zhou, S. Jarujaron, et al. (2007) HIV protease inhibitors increase TNF-alpha and IL-6 expression in macrophages: involvement of the RNA-binding protein HuR, Atherosclerosis. **195**, e134-143
59 C. Tudor, F. P. Marchese, et al. (2009) The p38 MAPK pathway inhibits tristetraprolin-directed decay of interleukin-10 and pro-inflammatory mediator mRNAs in murine macrophages, FEBS Lett. **583**, 1933-1938

60 L. E. Young, S. Sanduja, K. Bemis-Standoli, E. A. Pena, R. L. Price and D. A. Dixon. (2009) The mRNA binding proteins HuR and tristetraprolin regulate cyclooxygenase 2 expression during colon carcinogenesis, Gastroenterology. **136**, 1669-1679

61 B. Zhai, H. Yang, A. Mancini, Q. He, J. Antoniou and J. A. Di Battista. (2010) Leukotriene B(4) BLT receptor signaling regulates the level and stability of cyclooxygenase-2 (COX-2) mRNA through restricted activation of Ras/Raf/ERK/p42 AUF1 pathway, J Biol Chem. **285**, 23568-23580

62 A. Erol. (2007) Interleukin-6 (IL-6) is still the leading biomarker of the metabolic and aging related disorders, Med Hypotheses. **69**, 708

63 S. Omoigui. (2007) The Interleukin-6 inflammation pathway from cholesterol to aging--role of statins, bisphosphonates and plant polyphenols in aging and age-related diseases, Immun Ageing. **4**, 1

64 G. A. Tiberio, L. Tiberio, et al. (2007) Interleukin-6 sustains hepatic regeneration in cirrhotic rat, Hepatogastroenterology. **54**, 878-883

65 T. Kuilman, C. Michaloglou, L. C. Vredeveld, S. Douma, R. van Doorn, C. J. Desmet, L. A. Aarden, W. J. Mooi and D. S. Peeper. (2008) Oncogene-induced senescence relayed by an interleukin-dependent inflammatory network, Cell. **133**, 1019-1031

66 J. W. Kim, B. S. Baek, Y. K. Kim, J. T. Herlihy, Y. Ikeno, B. P. Yu and H. Y. Chung. (2001) Gene expression of cyclooxygenase in the aging heart, J Gerontol A Biol Sci Med Sci. **56**, B350-355

67 D. Wu, M. Marko, K. Claycombe, K. E. Paulson and S. N. Meydani. (2003) Ceramide-induced and age-associated increase in macrophage COX-2 expression is mediated through up-regulation of NF-kappa B activity, J Biol Chem. **278**, 10983-10992

68 E. H. Han, H. G. Kim, Y. P. Hwang, J. H. Choi, J. H. Im, B. Park, J. H. Yang, T. C. Jeong and H. G. Jeong. (2010) The role of cyclooxygenase-2-dependent signaling via cyclic AMP response element activation on aromatase up-regulation by o,p'-DDT in human breast cancer cells, Toxicol Lett. **198**, 331-341

69 K. Masuda, B. Marasa, J. L. Martindale, M. K. Halushka and M. Gorospe. (2009) Tissue- and age-dependent expression of RNA-binding proteins that influence mRNA turnover and translation, Aging (Albany NY). **1**, 681-698

70 K. Linker, A. Pautz, M. Fechir, T. Hubrich, J. Greeve and H. Kleinert. (2005) Involvement of KSRP in the post-transcriptional regulation of human iNOS expression-complex interplay of KSRP with TTP and HuR, Nucleic Acids Res. **33**, 4813-4827

71 A. Lal, K. Mazan-Mamczarz, T. Kawai, X. Yang, J. L. Martindale and M. Gorospe. (2004) Concurrent versus individual binding of HuR and AUF1 to common labile target mRNAs, Embo J. **23**, 3092-3102

72 N. Chang, J. Yi, et al. (2010) HuR uses AUF1 as a cofactor to promote p16INK4 mRNA decay, Mol Cell Biol. **30**, 3875-3886

CHAPTER II

ACF regulates the stability of IL-6 and Cox-2 mRNAs.

Abstract.

Based on the early embryonic lethality observed with germline *Acf* deletion and its widespread tissue distribution, we believe that ACF plays an extensive biological role beyond ApoB mRNA editing which needs to be further explored. ACF's cytoplasmic expression, shuttling activity, and homology to known AU-Rich RNA Binding Proteins (ARE-BPs), such as HuC, suggest that ACF may play a role in mRNA metabolism in the cytoplasm. The AU-rich nature of ACF's canonical binding site on ApoB mRNA [1] led us to hypothesize that ACF behaves as an ARE-BP that modulates mRNA stability and/or translation. Therefore, we undertook analysis of ACF's regulation of IL-6 and Cox-2, two AU-rich targets, in a variety of cell lines to determine ACF's role beyond ApoB mRNA editing.

Introduction.

Our previous work demonstrated that ACF acts as an ARE-BP that stabilizes IL-6 mRNA [2]. This work first began with the observation that partial hepatectomy in *Acf*$^{+/-}$ mice results in delayed liver regeneration when compared to wild type animals. These animals have fewer replicating liver cells (indicated by BrdU incorporation) and require nearly twice as long to recover liver mass. This observation was further examined to find that the delayed regeneration may be caused by decreased IL-6 mRNA levels. In wild type animals IL-6 is transcriptionally induced in Kupffer cells upon liver resection and secreted into the liver to induce cellular proliferation [3]. IL-6 appears to increase liver regeneration in a dose-dependent manner immediately following resection; plasma levels of IL-6 peak approximately 24 hours following partial hepatectomy [3, 4]. IL-6$^{-/-}$ mice eventually recover original liver mass, but on a significantly delayed time-scale, indicating that IL-6 is not absolutely essential for regeneration, but necessary for efficient re-growth. In *Acf*$^{+/-}$ mice, IL-6 mRNA is reduced at baseline, and though it is induced after liver resection, levels are significantly lower at points throughout the regeneration process. As IL-6's role in liver regeneration appears to be dose-dependent, it is possible that the ACF-dependent reduction of IL-6 mRNA (and thus serum IL-6) is involved in the delay of liver regeneration as this phenotype mimics that of IL-6 knockout or knockdown. However, it must be noted that if ACF regulates IL-6 as an ARE-BP, as postulated, it likely binds many other targets *in vivo* and *Acf*$^{+/-}$ mice may be exhibiting pleiotropic effects. It is also interesting to note that while IL-6 mRNA is reduced in *Acf*$^{+/-}$ mice, the transcript is induced upon liver resection. This indicates that ACF is not

impacting IL-6 expression on the transcriptional level, but likely regulates IL-6 abundance post-transcriptionally.

Further work established that isolated *Acf*$^{+/-}$ Kupffer cells show reduced half-life of endogenous IL-6 mRNA [2]. Together, these data suggest that ACF binds the IL-6 3'UTR and stabilizes IL-6 mRNA. This observation led us to consider a realm of other possible targets that ACF may modulate *in vivo*. In order to assess ACF's behavior as an ARE-BP, we examined its ability to regulate the stability of various RNA targets.

ACF binds and stabilizes IL-6 mRNA.

We began by extending our observations from Blanc et al. (2010) in which I and other authors used *in vitro* binding assays and RNA mutagenesis to demonstrate that ACF directly binds a fifteen nucleotide sequence in the 129-nt AU-Rich segment of the IL-6 3'UTR [2]. First, we addressed the *in vivo* ACF:IL-6 interaction. IL-6 is transcriptionally induced in SW480 colon adenocarcinoma cells following treatment with lipopolysacharide (LPS) (Figure 2.1A). From these treated cells, ACF was isolated in 100 mM KCl, 5 mM $MgCl_2$ from SW480 cells via αACF antibody bound to agarose beads. These conditions should be stringent enough in salt content to control for non-specific protein:RNA interactions occurring transiently. RNA was isolated from the pull-down material and assayed by RT-PCR using primers for the human IL-6 coding region. While ACF is predicted to bind the IL-6 3'UTR, the coding region should be pulled down as part of the ACF:IL-6 RNA complex. Its presence confirms the mRNA is still predominantly intact. ACF immunoprecipitation resulted in detectable IL-6 RNA while control IgG immunoprecipitation resulted in no detectable IL-6 RNA (Figure 2.1B). This

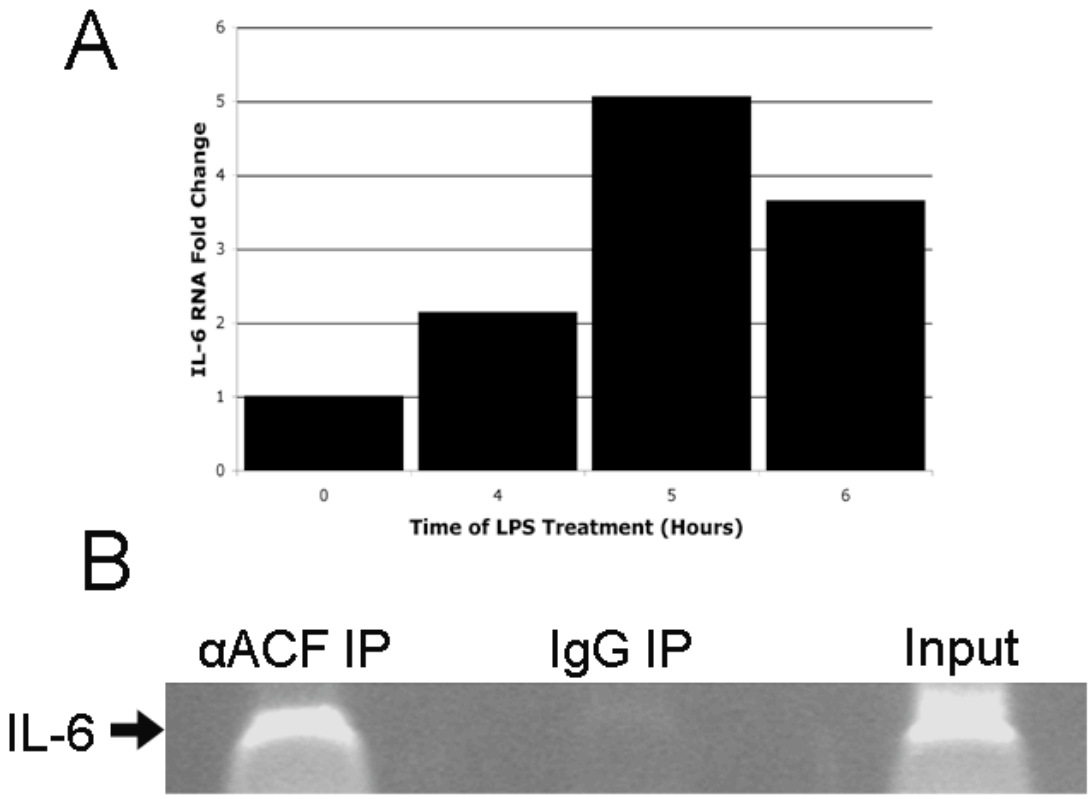


Figure 2.1 - ACF Binds IL-6 *in vivo*. A) IL-6 mRNA is induced in SW480 cells treated with LPS. Cells were treated with 2 μg/mL LPS. After the time indicated, total RNA was isolated with TRIZOL, cDNA generated (2 μg RNA), and assayed by qRT-PCR for IL-6. B) Cells treated for 6 hours with 2 μg/mL LPS were lysed under native conditions then exposed to αACF antibody or IgG bound to agarose beads. Beads were washed, proteinase treated, then RNA isolated by TRIZOL, cDNA generated (total sample), and assayed by RT-PCR for the human IL-6 coding region.

supports the co-immunoprecipitation data presented in Blanc et al. (2010) in which the RNA and protein were UV-crosslinked prior to immunoprecipitation. Our data suggest that the interaction is sufficiently robust to be detected without crosslinking, while the published UV-crosslinking immunoprecipitation data confirms this interaction occurs *in cellulo* and is not an artifact of the lysis process.

We next wanted to examine the effect of ACF on IL-6 stability. In order to compare the half-life of RNA in a cell line with normal levels of ACF (siControl) or one with significantly reduced ACF (siACF), we began by using siACF HepG2 hepatocellular carcinoma cells, where ACF is highly expressed. An siRNA targeted against ACF (siACF) or a random control siRNA (siControl) was stably transfected into HepG2 cells, resulting in 80% ACF knockdown (Figure 2.2A).

To begin, we wanted to assess the effect of ACF on the full IL-6 3'UTR. Therefore, we cloned the full length IL-6 3'UTR (Figure 2.2B) behind a Luciferase reporter coding region to create a construct we could track *in cellulo*. The Luciferase mRNA is transcribed under an SV40 promoter with the IL-6 3'UTR immediately downstream of the stop codon. This allows us to assay for Luciferase RNA as a biological readout of the impact of the IL-6 3'UTR in cell lines with varied expression of ACF. By comparing the half-life of this construct in siACF and siControl cell lines, we were able to observe the effect of ACF on IL-6 mRNA stability. Following actinomycin D treatment to halt transcription, the levels of Luciferase RNA were assessed by qRT-PCR. The Luciferase-IL-6 FL 3'UTR construct had a significantly reduced half-life in siACF cells compared to siControl cells (Figure 2.2C). These data were also replicated in a second siACF clone to confirm that the observed phenotype is not due to a clonal effect

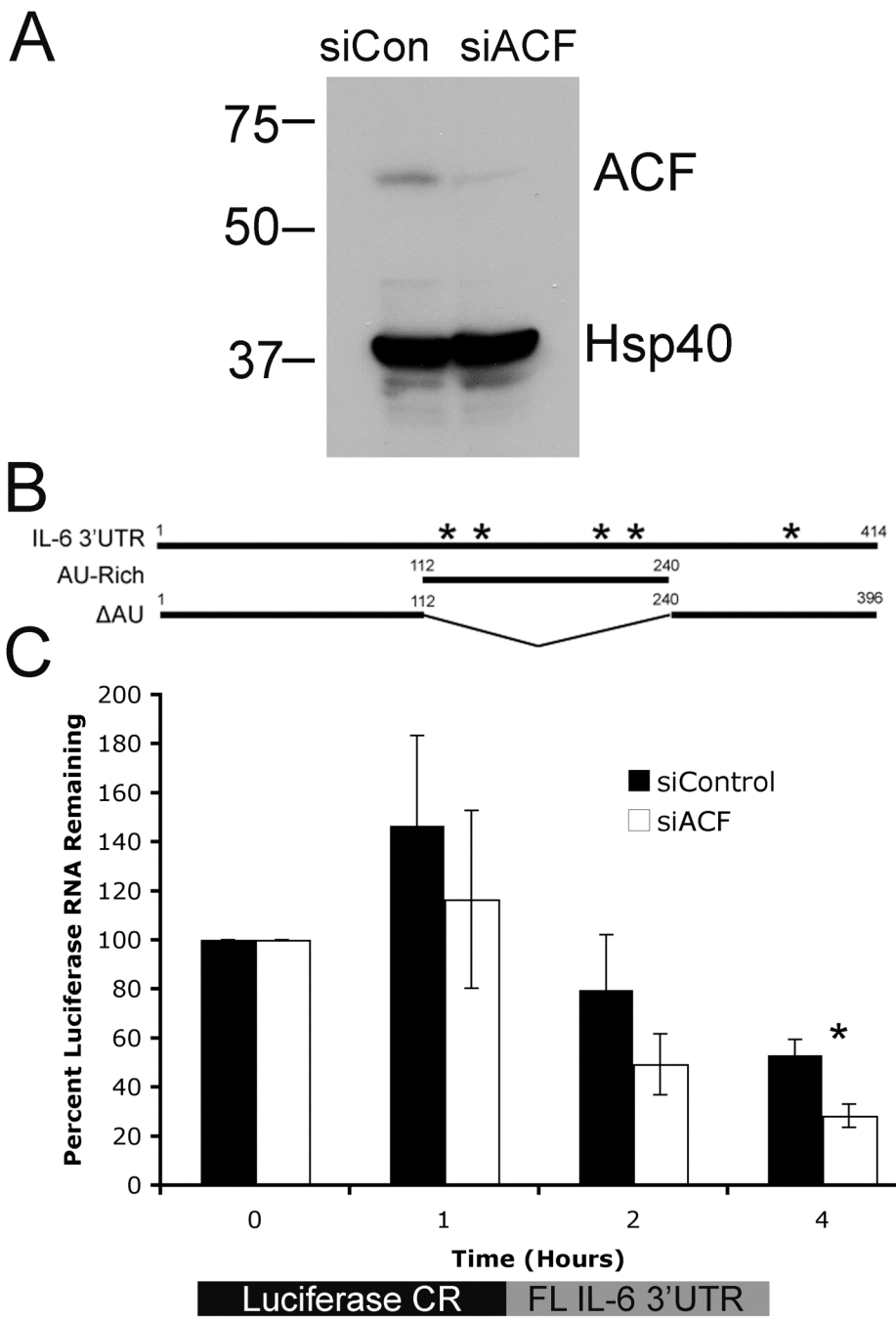


Figure 2.2 - ACF stabilizes Full Length IL-6 3'UTR. A) siControl (siCon) and siACF cell lysates were resolved by SDS-PAGE and western blotted with αACF antibody or αHsp40 antibody as a loading control. B) Schematic of IL-6 3'UTR. Numbering begins with first nucleotide of 3'UTR. * indicates AUUUA pentamer. C) siControl or siACF cells were transfected with the Luciferase-FL IL-6 3'UTR construct then treated with actinomycin D. After the indicated time, total RNA was collected, cDNA generated, and assayed for Luciferase RNA by qRT-PCR. n=3-5 ± standard error. *, $p < 0.04$

(data not shown), confirming that lower ACF expression results in reduced half-life of the reporter. This indicates that the higher expression of ACF in siControl cells has a stabilizing effect on the reporter through the IL-6 3'UTR.

To further explore the role of ACF in IL-6 mRNA stabilization, we probed truncation mutants of the IL-6 3'UTR to determine the region(s) involved in ACF-dependent stabilization. In order to narrow down the site of interaction, we cloned the 129 nt region of the IL-6 3'UTR (Figure 2.2B), which ACF has been reported to bind [2], behind the Luciferase coding region. We found that in control cells, the reporter gene was stable for at least 2 hours. However, in siACF cells the reporter gene was substantially less stable (Figure 2.3A). Once again, these data were replicated in a second siACF clone to control for clonal effects. In these cells the Luciferase-AU-rich reporter was longer-lived than observed in the original siACF cells, but still had a significantly reduced half-life compared to siControl cells (data not shown). Therefore, while the precise half-life appears to vary among clones, the trend of ACF-dependent stabilization remains constant. Together these data support our hypothesis that ACF's stabilization of IL-6 mRNA occurs via interaction with the 129 nt AU-rich region of the IL-6 3'UTR.

While this appears to be a robust stabilization phenotype upon expression of ACF, it is important to confirm that this is an ACF-specific event. Therefore, we used siACF HepG2 cells and transiently transfected them with either a control (empty) vector or an ACF-expression plasmid (under a CMV promoter) in order to rescue the siACF phenotype. We found that in cells expressing ACF at higher levels the Luciferase-IL-6 AU-rich 3'UTR construct is stable for at least two hours, similar to siControl cells. However, in control siACF cells (in which ACF expression remains reduced) Luciferase

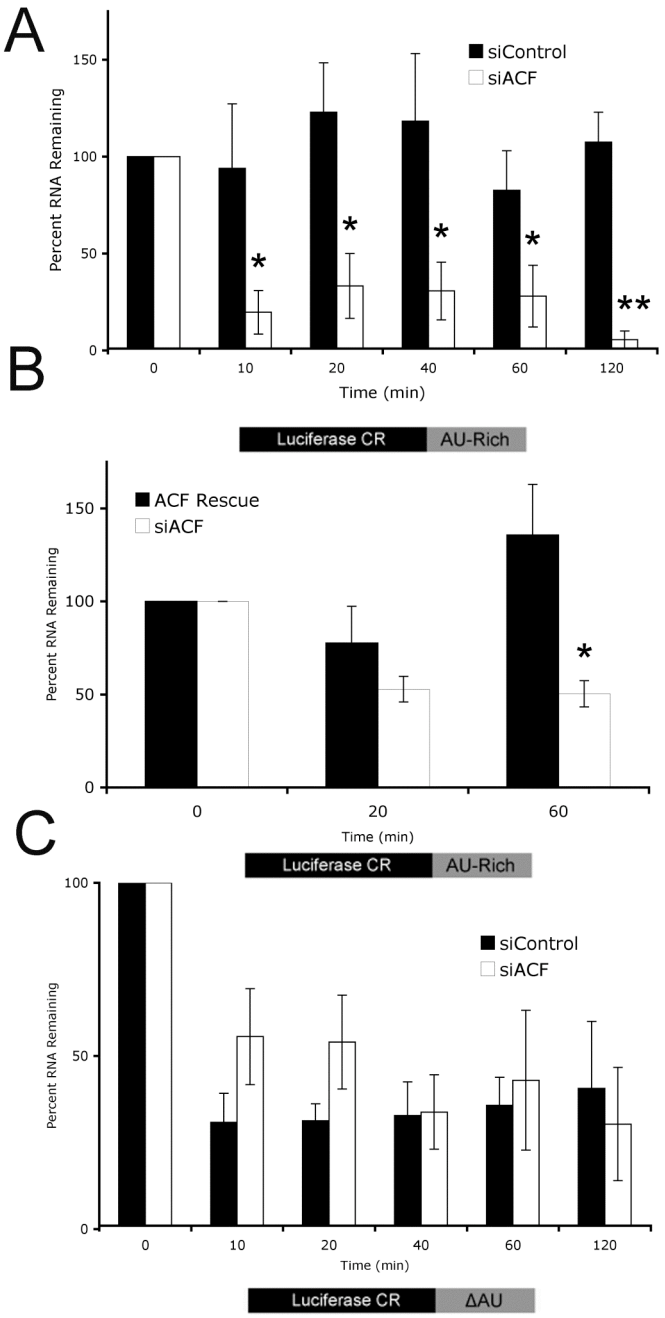


Figure 2.3 - ACF Stabilizes the AU-rich region of the IL-6 3'UTR. A) siControl or siACF cells were transfected with the Luciferase-AU-rich reporter construct and treated with actinomycin D for the times indicated. Total RNA was collected and qRT-PCR used to assay Luciferase RNA. B) siACF cells were transfected with a control or ACF-expression plasmid along with the Luciferase AU-rich reporter construct and half-life analysis conducted in the same manner as above. C) siACF or siCon cells were transfected with Luciferase-IL-6 3'UTR ΔAU reporter construct and half-life analysis conducted in the same manner as above. n=3-5 ± standard error. *, $p < 0.05$. **, $p < 0.01$

RNA has a significantly shorter half-life (Figure 2.3B). These data indicate that ACF add-back can rescue the phenotype observed in the siACF cells and supports the interpretation that ACF stabilizes IL-6 mRNA through interaction with this 129 nt region of the 3'UTR.

To further characterize this interaction, we created a Luciferase reporter construct in which the 129 nt "AU-Rich" region of the IL-6 3'UTR has been deleted (ΔAU, Figure 2.2B). The truncated UTR thus consists of nucleotides 1-111 and 240-396. The stability of this construct was once again examined in siACF and siControl cells. The half-life of Luciferase-IL-6 ΔAU was equally short in both cell lines, regardless of ACF expression (Figure 2.3C), suggesting that the ΔAU construct lacks the physiologically relevant ACF binding site, such that even when ACF is expressed, the protein cannot stabilize the RNA. These data further confirm our findings from Blanc et al (2010, [2]) and our above data that indicate ACF stabilizes IL-6 mRNA though interaction with the "AU-rich" region of the 3'UTR.

ACF destabilizes Cox-2 mRNA.

While the above data and previous work have shown that ACF stabilizes IL-6 mRNA, this property could be a general effect of ACF for any mRNA containing an ARE, or it could be unique to IL-6. To assess the generality of the observation, we characterized ACF's interaction with a novel target: Cox-2 mRNA. As discussed in the introductory chapter, Cox-2 is heavily regulated on the post-transcriptional level, particularly in the colon where ACF is highly expressed. The 3'UTR of Cox-2 is over 67% AU and the first 64 nucleotides contain an even higher concentration of A's and U's

and AUUUA pentamers. This region also includes the apparent HuR binding site [5], indicating it is important in the stability of Cox-2 mRNA. Therefore, this is the region in which we focused to explore ACF's interaction with and impact on Cox-2 mRNA.

We began by examining the effect of ACF expression on steady-state endogenous Cox-2 mRNA levels. HeLa cells express low levels of ACF. Therefore, we transfected HeLa cells with a control (empty) vector or a vector expressing ACF under a CMV promoter (Figure 2.4A), thus allowing us to examine the impact of ACF over-expression on endogenous Cox-2 mRNA levels. At baseline, ACF over-expressing cells exhibited a 50% reduction in endogenous Cox-2 mRNA compared to control cells (Figure 2.4B). This suggests that ACF leads to a significant reduction in Cox-2 mRNA though the mechanism could be transcriptional or post-transcriptional.

In an effort to more closely examine the mechanism by which ACF regulates Cox-2 mRNA expression, we used TNF-α treatment. In HeLa cells TNF-α transcriptionally induces Cox-2 mRNA expression through NF-κB activation of the Cox-2 promoter [6-8]. Therefore, we treated HeLa cells, transfected with a control (empty) or ACF-expression vector, with TNF-α. In both control and ACF over-expressing cells, Cox-2 mRNA was induced to a similar extent after TNF-α treatment. This suggests that both cell lines are sensitive to this particular cascade of Cox-2 transcriptional induction. However, while Cox-2 mRNA remained elevated at least six hours in control cells, in ACF over-expressing cells Cox-2 mRNA was reduced more than 50% by 4 hours after treatment (Figure 2.4C). These data suggest that ACF is not blocking transcriptional induction of Cox-2 but leads (directly or indirectly) to decreased mRNA levels through a different mechanism, possibly by post-transcriptional destabilization.

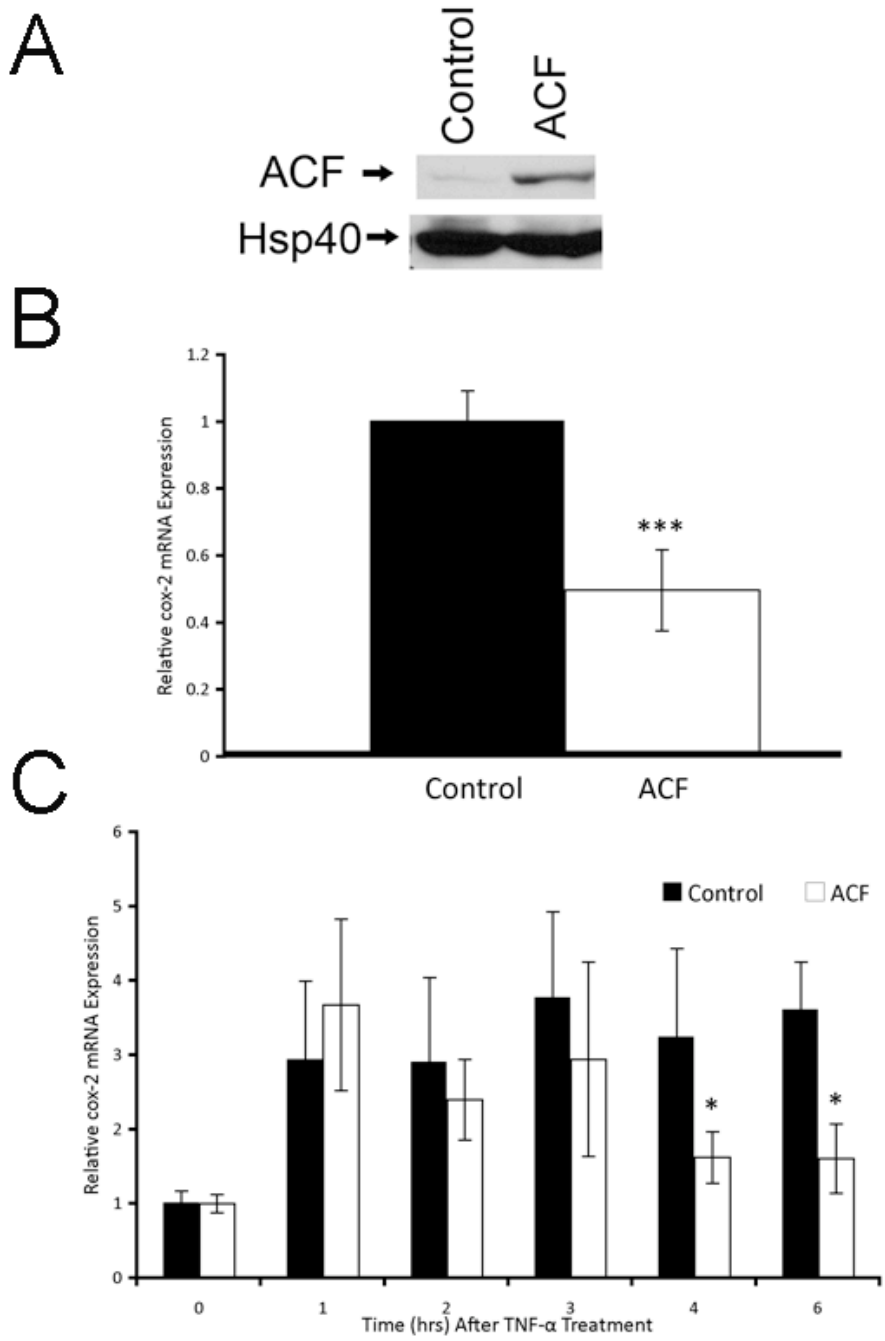


Figure 2.4 - ACF reduces the abundance of endogenous Cox-2 mRNA. A) 48 hours after HeLa cells were transfected with a control (empty) or ACF-expression vector, cells were lysed and western blotted for ACF expression. Hsp40 expression was assayed as a loading control. B) 48 hours after transfection total RNA was collected from transfected cells and assayed by qRT-PCR for Cox-2 abundance. C) HeLa cells transfected (for 48 hours) with control (empty) or ACF-expression vector were treated with 10 ng/mL TNF-α. At indicated times after treatment total RNA was collected and assayed by qRT-PCR for Cox-2 abundance. n=3-4 ± standard deviation. *, $p<0.05$. ***, $p<0.000002$.

To confirm our hypothesis that ACF regulates Cox-2 mRNA stability, we undertook actinomycin D assays. We began with HepG2-Tet-On cells, in which over-expression of ACF is conditionally induced by Doxycycline treatment. These cells express ACF at baseline, but when treated with Doxycycline, ACF expression increases robustly (Figure 2.5A). Under these conditions we were able to observe the effect of varied levels ACF on a reporter construct containing a portion of the Cox-2 3'UTR.

Initially, we assayed a reporter mRNA that contained the first 64 nt of the Cox-2 3'UTR following the Luciferase coding region (Figure 2.5B). This segment of the 3'UTR has the highest concentration of AU-rich sequences and is the most likely site of ACF interaction. In control cells (-Dox), Luciferase mRNA was stable for at least 4 hours following actinomycin D treatment when assayed by qRT-PCR. However, in cells over-expressing ACF (+Dox), the half-life of the reporter construct was reduced to approximately 3.5 hours (Figure 2.5C). This is consistent with the TNF-α induction Cox-2 data presented above, suggesting ACF destabilizes Cox-2 mRNA with a $t_{1/2}$ in the range of 3-4 hours (Figure 2.2C). To further confirm the effect of ACF on Cox-2 mRNA, we mutated a portion of Cox-2 3'UTR nt 1-64 to eliminate a potential ACF-interacting site by substituting C residues for U residues from nt 37-46 (Figure 2.5C). Unlike the wild-type sequence, this mutant reporter failed to exhibit ACF-induced destabilization within the four-hour experimental course (Figure 2.5D), indicating we had mutated at least part of the ACF interaction site.

To confirm these observations are not unique to the HepG2 Tet-On cell line, we conducted similar experiments in HEK 293 cells (Human embryonic kidney). Much like HeLa cells, HEKs express low levels of ACF at steady state. Therefore, cells were

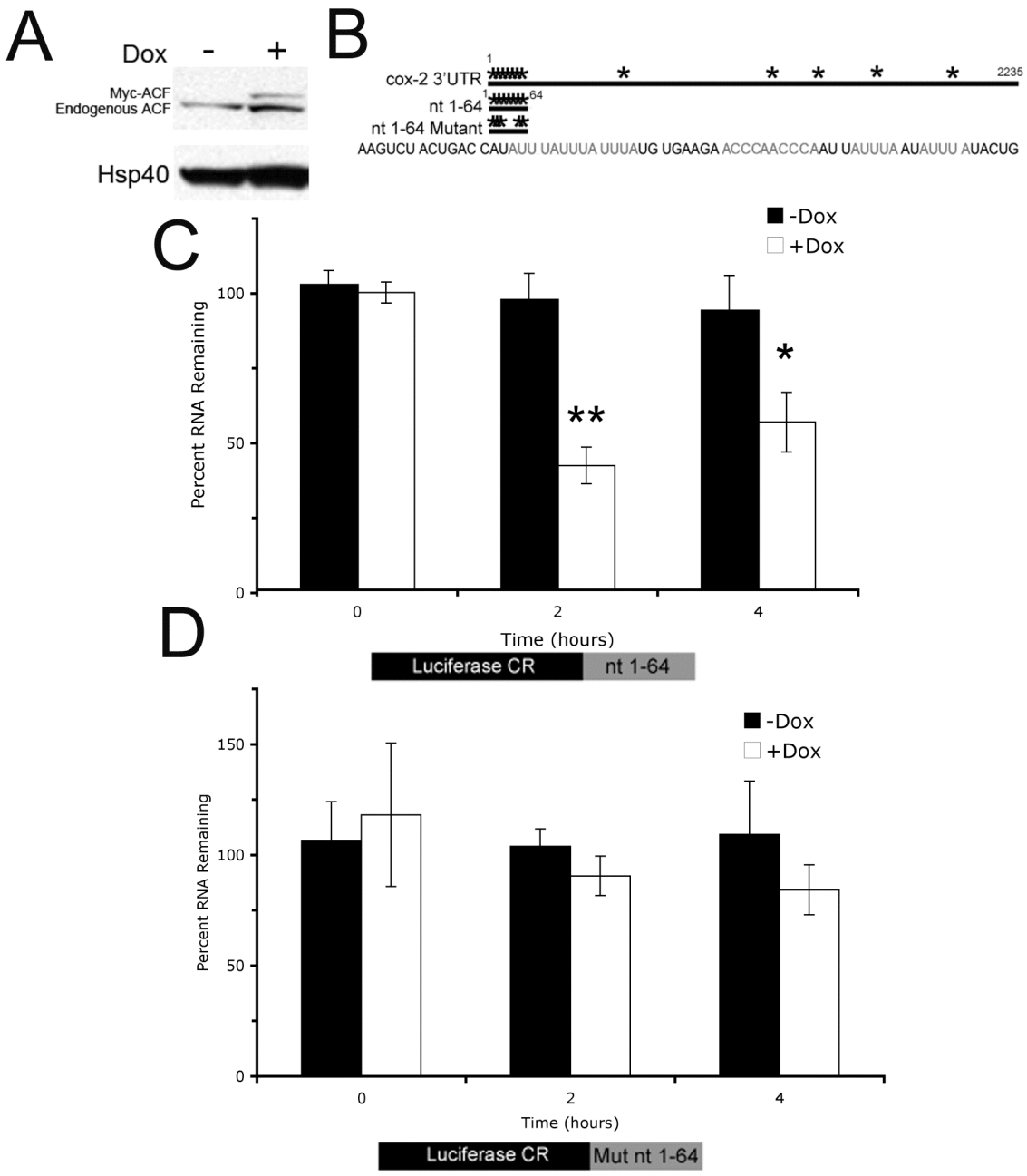


Figure 2.5 - ACF destabilizes the Cox-2 3'UTR. A) ACF over-expression is induced by Doxycycline treatment in HepG2 Tet-On cells. B). Schematic of the Cox-2 3'UTR. Numbering begins with the first nucleotide of the 3'UTR. * Indicates AUUUA pentamer. Primary sequence of "Mutant nt 1-64" is shown. C) Luciferase-Cox-2 nt 1-64 or D) Luciferase Cox-2 Mutant nt 1-64 was transfected into HepG2 Tet-On cells with or without Doxycylcine treatment to induce ACF over-expression. Cells were treated with actinomycin D and after times indicated total RNA was collected and assayed for Luciferase RNA by qRT-PCR. n = 3-6 ± standard error. *, $p < 0.04$. **, $p < 0.001$.

transfected with a control (empty) or ACF-expressing vector along with the Luciferase-Cox-2 reporter constructs employed above. The Luciferase-Cox-2 nt 1-64 wild type reporter was stable in control cells but in ACF over-expressing cells turned over more rapidly (Figure 2.6A). However, the Lucerifase-Cox-2 nt 1-64 mutant construct failed to turn over in either cell line during the one hour experimental timeframe (Figure 2.6B). It should be noted that the apparent half-life of the reporters is different in HepG2 Tet-On cells compared to HEK 293 cells. This is likely the result of varied expression levels of other Cox-2 regulatory ARE-BPs such as HuR that may be modulating the half-life of the reporter independent of ACF or interacting with ACF to alter its affect. However, these HEK293 data support the above observations in the HepG2 Tet-On cells to confirm that this phenotype trend is not unique to a single cell line.

The Cox-2 reporter turnover data indicate that ACF interacts with the Cox-2 3'UTR, specifically via nucleotides 37-46, to destabilize the transcript. In concert with the earlier data which suggest ACF results in decreased endogenous Cox-2 mRNA abundance through a post-transcriptional mechanism, we predict that ACF destabilizes Cox-2 mRNA *in vivo* by binding the Cox-2 3'UTR.

ACF interacts with other ARE-BPs.

Our data above clearly indicate that ACF regulates the mRNA stability of at least two targets: IL-6 and Cox-2. However, these two RNAs are regulated by a host of other ARE-BPs, specifically, HuR and Auf1. Both of these proteins contain RRMs and shuttle from the nucleus to the cytoplasm where they regulate mRNA stability [9]. HuR and Auf1 have been observed to both homodimerize and heterodimerize with each other, in

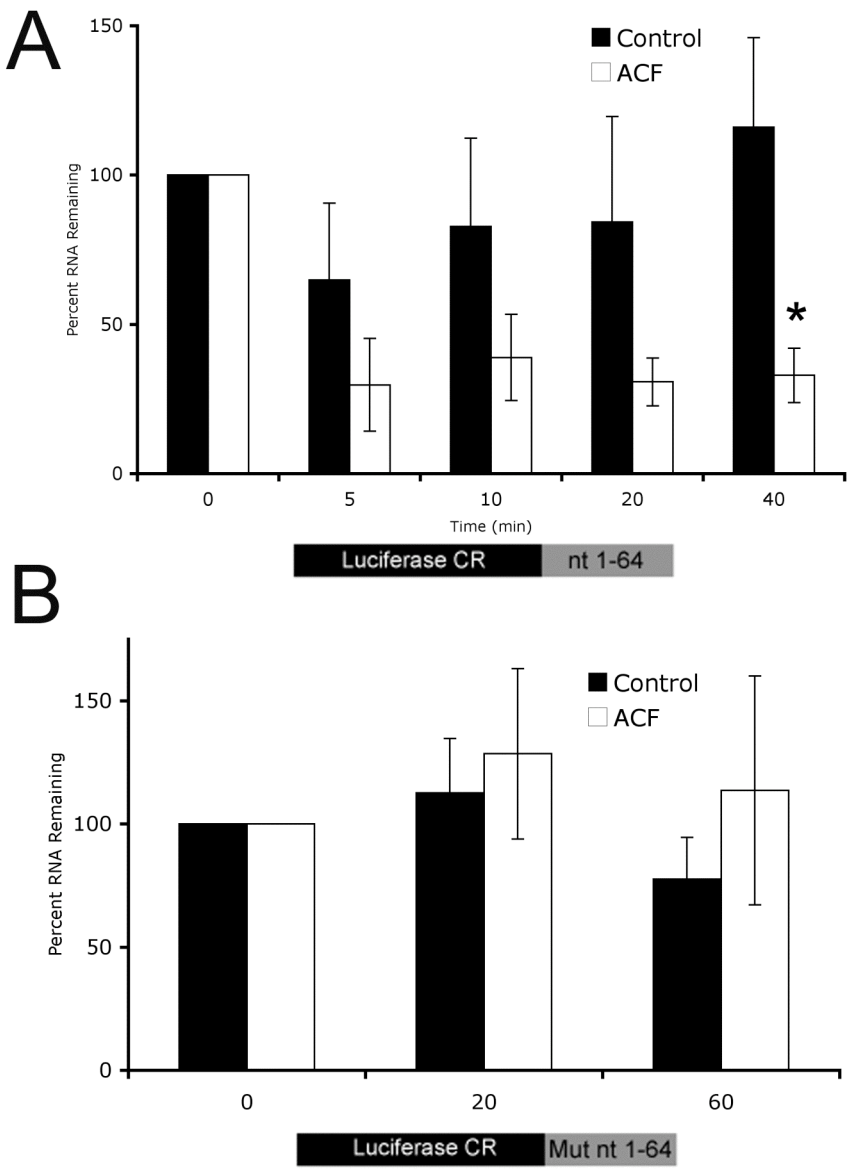


Figure 2.6 - ACF destabilizes Cox-2 3'UTR in HEK cells. HEK 293 cells were transfected with a control (empty) or ACF-expression vector along with (A) Luciferase-Cox-2 nt 1-64 wild type or (B) Luciferase-Cox-2 Mutant nt 1-64. 48 hours after transfection, transcription was halted with actinomycin D and total RNA collected at indicated times after treatment and assayed for Luciferase RNA by qRT-PCR. n=3-6 ± standard error. *, $p < 0.03$.

an RNA-dependent manner [9-11]. This led us to question if ACF interacts with either of these ARE-BPs.

We began by examining the intercellular localization of each protein. In HepG2 cells we were able to detect endogenous ACF (using primary rabbit αACF IgG and secondary αrabbit IgG conjugated to FITC, which excites green) and endogenous HuR (using primary mouse αHuR IgG and secondary αmouse IgG conjugated to Cy3, which excites red) (Figure 2.7A). Nuclei were detected by DAPI staining (shown in blue). Both ACF and HuR appear to localize to the nucleus, (indicated by the yellow signal in the red + green merge), though ACF is detectable in the cytoplasm as well, particularly noticeable in the red halos around the areas of co-expression in the merged image. It is important to note that these images are not from confocal microscopy; therefore we are visualizing regions of co-expression, but not definite co-localization.

Due to antibody limitations (both the αACF and αAuf1 antibodies are rabbit IgG) we were unable to detect co-expression of endogenous ACF and Auf1; therefore, HEK293 cells were co-transfected with Flag-ACF (detected by primary rabbit αFLAG IgG and secondary αrabbit IgG conjugated to FITC, which excites green) and myc-Auf1 (detected by primary mouse α-Myc IgG and secondary αmouse IgG conjugated to Cy3, which excites red) (Figure 2.7A). Once again, nuclei were identified by DAPI staining (shown in blue). Interestingly, transfected Flag-ACF appears to localize entirely to the nucleus, unlike endogenous ACF, which is expressed to a small extent in the cytoplasm. However, these tagged forms of ACF and Auf1 both predominantly localize to the nucleus. These data suggest that ACF is expressed in the same subcelluar compartments and therefore in close proximity to both HuR and Auf1 *in vivo.*

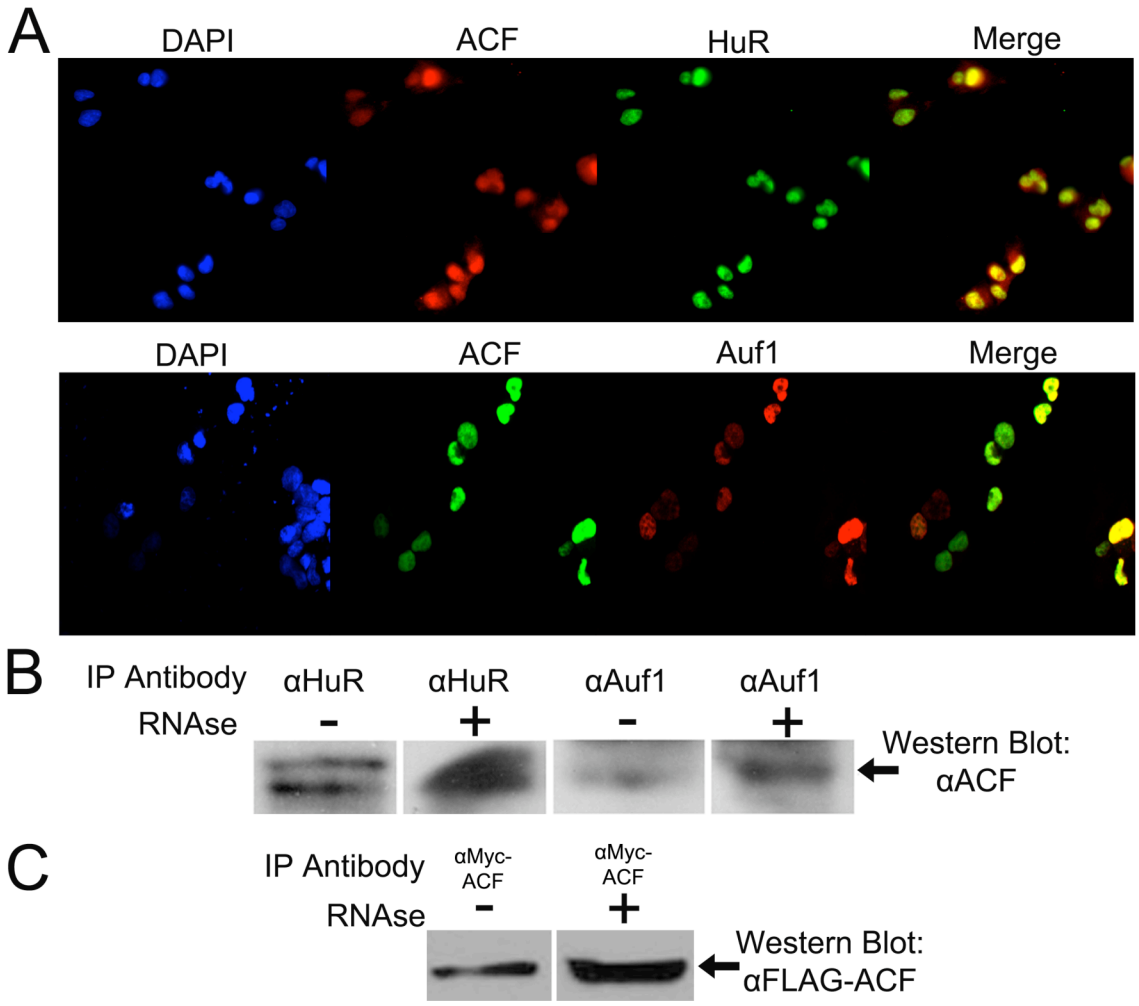


Figure 2.7 - ACF interacts with HuR and Auf1. A) Top: Endogenous ACF (red) and HuR (green) in HepG2 cells. Bottom: Flag-ACF (green) and Myc-Auf1 (red) in HEK 293 cells. B) Co-immunoprecipitation of ACF:HuR or ACF:Auf1 (in 50 mM Tris pH 8.0, 100 mM NaCl, 0.5% Triton X-100, 1 mM DTT, 1 mM EDTA, protease inhibitor cocktail) in the presence and absence of RNase from HepG2 cells. C) Co-immunoprecipitation of myc-ACF and Flag-ACF from co-transfected Cos-7 cells in the presence and absence of RNase.

In order to more directly assess the interaction between these ARE-BPs, we turned to protein co-immunoprecipitation. Endogenous HuR and Auf1 were immunoprecipitated from HepG2 cells (in 50 mM Tris pH 8.0, 100 mM NaCl, 0.5% Triton X-100, 1 mM DTT, 1 mM EDTA) with specific antibodies bound to Protein A-agarose beads and the immunoprecipitate was western blotted for ACF expression. ACF was pulled-down along with both HuR and Auf1, indicating these proteins are capable of interaction (Figure 2.7B). However, this may be a direct protein:protein interaction or alternatively be due to co-localization to the same RNA, as the case for HuR:Auf1 [9, 12]. Therefore, we performed the same co-IP experiments in the presence of RNase A and RNase T1 in order to eliminate any RNA that may be linking the proteins. In the presence of RNase, ACF was pulled down with both HuR and Auf1. This indicates that ACF is capable of directly interacting with both of these ARE-BPs independent of RNA binding.

It is also interesting to note that in our co-immunoprecipitation experiments, we observed ACF homodimerization. A truncated form of ACF (the first 320 amino acids) has been shown to self-associate in an RNA-dependent manner [13] via size-exclusion chromatography and dynamic light scatter assays. However, here we show that full length ACF can homodimerize in an RNA-independent manner. Cos-7 cells were co-transfected with myc-ACF and Flag-ACF. Cells were then lysed in the presence or absence of RNase A and RNase T1. α-Myc antibody was used to immunoprecipitate myc-ACF and product was probed by α-FLAG antibody via western blot to determine the presence of Flag-ACF. As indicated in Figure 2.7C, myc-ACF and Flag-ACF co-immunoprecipitated both

in the presence or absence of RNase, suggesting full length ACF is capable of homodimerization via direct protein-protein interaction.

Conclusion.

In light of published observations as well as our own data, we presume that ACF functions as an ARE-BP in the liver to stabilize IL-6 mRNA post-transcriptionally. Due to ACF's RNA binding capabilities and homology to numerous ARE-BPs, it is not surprising to find that ACF's role beyond ApoB mRNA editing lies in mRNA stability regulation through AREs. However, due to the convincing evidence that ACF stabilizes IL-6 mRNA, it was relatively unexpected to find that ACF destabilizes Cox-2 mRNA. This introduces an interesting characteristic of ACF as an ARE-BP, as it is able to interact with multiple RNAs to differentially regulate their stability. This is not a particularly unique behavior, as HuR has traditionally been observed to stabilize targets but has recently been implicated in the destabilization of some mRNAs. However, to date there has been no explanation for these proteins' abilities to bind multiple targets and differentiate which should be stabilized and which destabilized.

It is also interesting to note that ACF is capable of directly interacting with both itself as well as HuR and Auf1, which have been observed to regulate the stability of numerous mRNAs, including IL-6 and Cox-2. It is highly likely that these RNAs are regulated by a complement of proteins *in vivo,* which work in concert or competition to result in the modulation of IL-6 and Cox-2 gene product expression. However, it is difficult to predict how these higher-order interactions may result in mRNA stabilization or destabilization, particularly given the little information known about the manner by

which ACF interacts with its mRNA targets. We propose the biochemical mechanism of RNA binding may contribute to these protein's abilities to seek their targets in concert and modulate their stability in response to environmental stimuli.

MATERIALS AND METHODS

LPS Treatment and RNA Co-Immunoprecipitation – SW480 cells were treated with 2 μg/mL LPS for times indicated. Total RNA was collected via TRIZOL reagent and assayed by qRT-PCR for human IL-6, normalized against 18S RNA. For co-immunoprecipitation, SW480 cells were treated for 6 hours with 2 μg/mL LPS then collected in 10 mM Hepes pH 7.0, 100 mM KCl, 5 mM $MgCl_2$, 0.5% NP-40, 1 mM DTT, 100 U/mL RNaseOut, 0.2% Vanadyl Ribonucleoside Complexes, 0.2 mM PMSF, Protease Inhibitor Cocktail (Roche) and lysed by sonication. Lysate was cleared twice by centrifugation and assayed for protein concentration. Equivalent amounts of cleared lysate were added to each of two preparations of Protein-A agarose beads linked to αACF or unrelated rabbit IgG antibodies at 4°C overnight. Beads were washed five times in 50 mM Tris pH 7.4, 150 mM NaCl, 1 mM $MgCl_2$, 0.05% NP-40. Washed beads were pelleted, proteinase K treated, and exposed to TRIZOL reagent. Total RNA was isolated and used to generate cDNA with random hexamer primers. cDNA was exposed to PCR with primers for human the IL-6 coding region. PCR was product resolved on 1% agarose gel.

Bead prep: Protein A-agarose beads (Promega) were pre-swollen in 50 mM Tris pH 7.5, 150 mM NaCl, 1 mM $MgCl_2$, 0.05% NP-40 then incubated overnight at 4°C with 25 μg rabbit polyclonal αACF or unrelated rabbit IgG antibody. Beads were then washed three times in buffer prior to use.

RNA extraction and qRT-PCR – Total RNA was collected using TRIZOL reagent. Q-PCR was performed on reverse-transcribed DNase-treated RNA (2 μg) with an ABI Prism 7000 instrument (Applied Biosystems) using SYBR Green Master Mix according to the manufacturer's instructions. Luciferase-IL-6 reporter abundance was normalized against 18S RNA. Endogenous Cox-2 abundance was normalized against GAPDH RNA. Luciferase-Cox-2 reporter abundance was normalized against expression of a Neomycin-Resistance element on the co-transfected pTRE-Myc-ACF plasmid.

ACF protein expression – Cells as indicated were collected in 20 mM Tris pH 7.5, 1 mM Sodium Vanadate, 150 mM NaCl, 2 mM EDTA, 100 mM NaF, 5% Glycerol, 50 mM B-glycerophosphate, 10% Triton X-100, 1% SDS, Protease Inhibitor Cocktail (Roche), lysed by syringe and cleared by centrifugation. Protein concentration determined by Biorad Protein Assay and 40 μg total protein resolved by SDS-PAGE and western blotted with αACF or αHsp40 antibody (Stressgen).

Cloning of Reporter Constructs: IL-6 reporter constructs were PCR amplified from previously described plasmids [2] and inserted into pGL3 in the XbaI site. The wild-type Cox-2 nt 1-64 construct in pGL3 was a generous gift from Dr. Aubrey Morrison (Washington University in St. Louis). For the Cox-2 mutant, top and bottom oligonucleotides were annealed with the appropriate cohesive ends and ligated into pGL3 in the XbaI/FseI sites.

IL-6 mRNA stability assays – Stably siRNA transfected siACF or siControl cells were transiently transfected with 2 μg Luciferase reporter construct via FuGene reagent (Qiagen) per manufacturers instructions. 48 hours after transfection, cells were treated with 5 μg/mL actinomycin D (Sigma) and collected in TRIZOL reagent at indicated

times. In ACF rescue experiment, siACF cells were transfected (FuGene, Qiagen) with 1 μg Luciferase-AU-rich reporter plasmid and 1 μg pCMV2B (control) or pCMV2B-ACF (cloning described in [14]).

Endogenous Cox-2 assays – HeLa cervical cancer cells were transfected with 2 μg pCMV2B-ACF or pCMV2B (control) with FuGene reagent (Qiagen) and checked for ACF protein expression via western blot. 48 hours after transfection, cells were treated with media supplemented with 10 ng/mL recombinant TNF-α (Sigma) and collected in TRIZOL reagent at indicated times. For steady state experiments samples from ACF cells were reported relative to control cell levels. For TNF-α induction experiments, all control cells were compared to the control 0 hour samples and all ACF cells were compared to the ACF 0 hour samples.

Cox-2 mRNA stability assays – Tet-On HepG2 cells were stably transfected with pN1pbactin-rtTA2S-M2-IRES-EGFP (a generous gift from Dr. Arkadiusz Welman, University of Edinburgh) [15]. For construction of pTRE-Myc-ACF, human ACF was cloned into pTRE2pur-myc plasmid (Clontech) into NotI and EcoRI restriction sites, resulting in the doxycycline-dependent expression of an N-terminal myc-tagged ACF. HepG2 clones stably expressing pN1pβactin-rtTA2S-M2-IRES-EGFP were transiently transfected with 1 μg pTRE-Myc-ACF and 1 μg Luciferase reporter vector with FuGene reagent (Qiagen). At 48 hours, "+Dox" cells were treated with 2 μg/mL Doxycycline to induce ACF overexpression. 72 hours after transfection, all cells were treated with 10 μg/mL actinomycin D and collected in TRIZOL reagent at the indicated times. HEK cells were transfected with 1 μg Luciferase reporter and 1 μg pCMV2B or pCMV2B-ACF and

48 hours later treated with 10 μg/mL actinomycin D and collected in Trizol reagent at the indicated times.

Protein Localization – ACF/HuR co-staining: HepG2 cells were grown to 50-60% confluency on coverslips. Cells were fixed with 10% formalin solution, permeabilized with 0.5% Triton X-100, and probed with rabbit polyclonal αACF IgG and mouse monoclonal αHuR IgG (Santa Cruz) followed by Cy3 secondary IgG and fluorescein isothiocyanate (FITC)-conjugated secondary IgG (Jackson ImmunoResearch). Nuclei were identified using 4,6-diamidino-2-phenylindole (DAPI, Vector).

ACF/Auf1 co-staining: HEK 293 cells were grown on coverslips and co-transfected with 1 μg myc-Auf1 and 1 μg flag-ACF with FuGENE reagent. 48 hours after transfection cells were fixed with 10% formalin solution, permabilized with 0.5% Triton X-100, and probed with rabbit α-FLAG IgG (Jackson Immuno Research) and mouse monoclonal α-Myc IgG (Santa Cruz) followed by Cy3 secondary IgG and FITC-conjugated secondary IgG (Jackson Immuno Research). Nuclei were identified using DAPI stain (Vector).

All slides were examined using a Zeiss Axioskop 2 MOT microscope equipped with a 40x plan neofluar objective and a 3CCR camera (DAGE-MTI, Inc.) A Zeiss Attoarc variable intensity lamp was used with filter sets designed for Cy3, FITC, and DAPI. Images were processed using Adobe Photoshop software.

Protein Co-Immunoprecipitation – HuR and Auf1 Co-IP: HepG2 cells were collected in 50 mM Tris pH 8.0, 100 mM NaCl, 0.5% Triton X-100, 1 mM DTT, 1 mM EDTA, protease inhibitor cocktail and lysed by syringe then cleared by centrifugation. RNase treated samples were then exposed to 1 μg RNase A and 10 U RNase T1 for 10 minutes at 30°C. Lysate was then supplemented with 2 μg mouse monoclonal αHuR IgG (Santa

Cruz) or rabbit polyclonal αAuf1 (Upstate) at 4°C for 2 hours. 50 μL Protein A-agarose beads (Promega, pre-washed in lysis buffer) were added to each reaction at 4°C overnight. Beads were washed five times with lysis buffer, then boiled in denaturing protein loading buffer, resolved by 10% SDS-PAGE and western blotted with rabbit polyclonal αACF IgG.

ACF/ACF Co-IP: Cos-7 cells were co-transfected via FuGene reagent with 1 μg myc-ACF and 1 μg Flag-ACF expression vectors. 48 hours after transfection, cells were lysed by syringe in 50 mM Tris pH 8.0, 100 mM NaCl, 0.5% Triton X-100, 1 mM DTT, 1 mM EDTA, protease inhibitor cocktail then cleared by centrifugation. RNase treated samples were exposed to 40 μg RNase A and 1000 U RNase T1 for 30 minutes at 30°C. Lysate was then supplemented with 2 μg mouse α-Myc IgG at 4°C for 2 hours. 50 μL Protein A-agarose beads (pre-washed in lysis buffer) were added to each reaction at 4°C overnight. Beads were washed five times with lysis buffer, then boiled in denaturing protein loading buffer, resolved by 10% SDS-PAGE and western blotted with rabbit α-FLAG IgG.

Statistics – All data are reported as mean ± standard error. Statistical significance (p-values) was determined by a Student's 2-tailed T-test.

REFERENCES

1 V. Blanc, N. Navaratnam, J. O. Henderson, S. Anant, S. Kennedy, A. Jarmuz, J. Scott and N. O. Davidson. (2001) Identification of GRY-RBP as an apolipoprotein B RNA-binding protein that interacts with both apobec-1 and apobec-1 complementation factor to modulate C to U editing, J Biol Chem. **276**, 10272-10283

2 V. Blanc, K. J. Sessa, S. Kennedy, J. Luo and N. O. Davidson. (2010) Apobec-1 complementation factor modulates liver regeneration by post-transcriptional regulation of interleukin-6 mRNA stability, J Biol Chem. **285**, 19184-19192

3 G. K. Michalopoulos and M. C. DeFrances. (1997) Liver regeneration, Science. **276**, 60-66

4 G. A. Tiberio, L. Tiberio, et al. (2007) Interleukin-6 sustains hepatic regeneration in cirrhotic rat, Hepatogastroenterology. **54**, 878-883

5 S. J. Cok, S. J. Acton and A. R. Morrison. (2003) The proximal region of the 3'-untranslated region of cyclooxygenase-2 is recognized by a multimeric protein complex containing HuR, TIA-1, TIAR, and the heterogeneous nuclear ribonucleoprotein U, J Biol Chem. **278**, 36157-36162

6 G. Totzke, K. Schulze-Osthoff and R. U. Janicke. (2003) Cyclooxygenase-2 (COX-2) inhibitors sensitize tumor cells specifically to death receptor-induced apoptosis independently of COX-2 inhibition, Oncogene. **22**, 8021-8030

7 B. Kaltschmidt, R. A. Linker, J. Deng and C. Kaltschmidt. (2002) Cyclooxygenase-2 is a neuronal target gene of NF-kappaB, BMC Mol Biol. **3**, 16

8 S. Shishodia, D. Koul and B. B. Aggarwal. (2004) Cyclooxygenase (COX)-2 inhibitor celecoxib abrogates TNF-induced NF-kappa B activation through inhibition of activation of I kappa B alpha kinase and Akt in human non-small cell lung carcinoma: correlation with suppression of COX-2 synthesis, J Immunol. **173**, 2011-2022

9 P. S. David, R. Tanveer and J. D. Port. (2007) FRET-detectable interactions between the ARE binding proteins, HuR and p37AUF1, RNA. **13**, 1453-1468

10 N. Chang, J. Yi, et al. (2010) HuR uses AUF1 as a cofactor to promote p16INK4 mRNA decay, Mol Cell Biol. **30**, 3875-3886

11 G. M. Wilson, Y. Sun, H. Lu and G. Brewer. (1999) Assembly of AUF1 oligomers on U-rich RNA targets by sequential dimer association, J Biol Chem. **274**, 33374-33381

12 A. Lal, K. Mazan-Mamczarz, T. Kawai, X. Yang, J. L. Martindale and M. Gorospe. (2004) Concurrent versus individual binding of HuR and AUF1 to common labile target mRNAs, Embo J. **23**, 3092-3102

13 C. A. Galloway, A. Kumar, J. Krucinska and H. C. Smith. (2010) APOBEC-1 complementation factor (ACF) forms RNA-dependent multimers, Biochem Biophys Res Commun. **398**, 38-43

14 V. Blanc, J. O. Henderson, S. Kennedy and N. O. Davidson. (2001) Mutagenesis of apobec-1 complementation factor reveals distinct domains that modulate RNA binding, protein-protein interaction with apobec-1, and complementation of C to U RNA-editing activity, J Biol Chem. **276**, 46386-46393

15 A. Welman, J. Barraclough and C. Dive. (2006) Generation of cells expressing improved doxycycline-regulated reverse transcriptional transactivator rtTA2S-M2, Nat Protoc. **1**, 803-811

CHAPTER III

ACF binds IL-6 and Cox-2 RNAs *in vitro.*

Abstract.

While previous work has narrowed down the regions of ACF that interact with Apobec-1 and ApoB mRNA [3], these analyses were not comprehensive. The contributions of individual RRMs have not been fully examined, nor has it been established if ACF binds all potential RNA targets in the same manner. Because three physiological ACF:RNA targets of different sequence have been identified, it is also of interest to study the range of ACF's binding preferences. Therefore, to investigate the mechanism of ACF interaction with RNA, we undertook *in vitro* RNA binding studies. *In vitro* binding studies are useful to examine the interaction between a protein and its target RNA(s) under controlled conditions. Using purified protein with *in vitro* transcribed RNA, we compared ACF's affinity for targets of different sequence and structure. In these binding studies, we used a panel of ACF constructs (Figure 3.1A); full length ACF was informative in probing specific binding sites and truncation mutants were used to examine the domains of ACF that interact with various RNA targets.

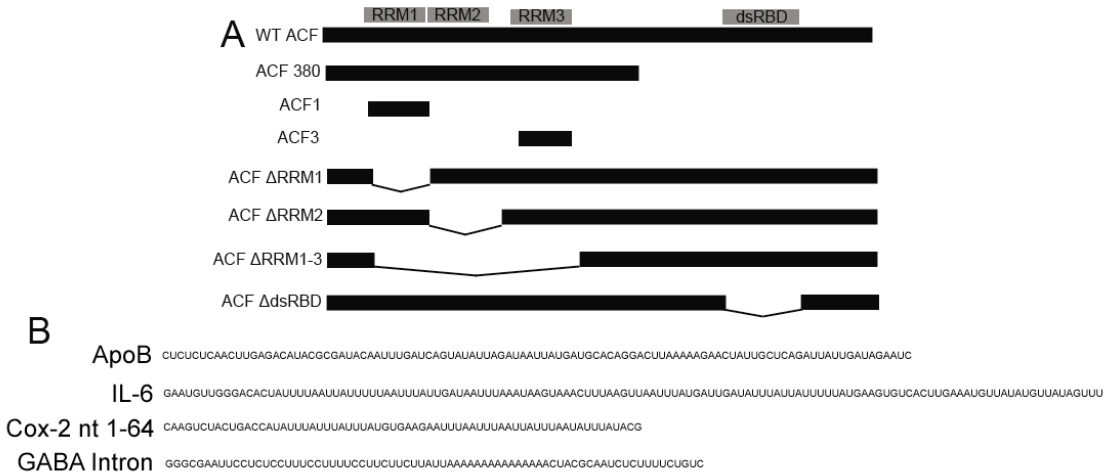


Figure 3.1 - Constructs used in *in vitro* binding experiments. A) Purified recombinant ACF constructs with putative domains noted at top. B) Primary RNA sequences used in EMSAs.

Construction of ACF Domains.

To facilitate examination of ACF:RNA interactions, three protein constructs were generated: ACF380, ACF1 and ACF3 (Figure 3.1A). ACF380 contains the first 380 amino acids of ACF, consistent with the ACF 43 and ACF 45 splice variants capable of binding ApoB mRNA and interacting with Apobec-1 [4]. ACF1 extends from amino acid 58 to 134, and ACF3 from 231 to 303. The structure of our ACF1 construct was modeled with the SWISS-MODEL homology-based structure prediction algorithm [5, 6] and is predicted to be consistent with the βαββαβ secondary structure and the α/β sandwich tertiary structure of an RRM (data not shown). The structure of ACF3 has been determined by solution NMR [7], and it adopts the canonical RRM topology. For our experiments, folding of our purified ACF380, ACF1 and ACF3 proteins was characterized by far-UV circular dichroism spectroscopy. As indicated in Figure 3.2A, all three proteins exhibit CD spectra consistent with a protein containing a mixture of α-helices and β-strands. We assume that the tertiary structure of each is correct.

RNA Binding.

Electrophoretic mobility shift assays were used to examine ACF380's RNA binding activity. While filter binding is a more sensitive and quantitative technique, it can be challenging with RNAs as long as those targeted by ACF (Figure 3.1B). Also, while ACF has been observed to dimerize both in the literature and our own observations (Chapter 2) [8], we do not know the stoichiometry of ACF:RNA complexes, therefore filter binding data would be un-interpretable. Therefore, we turned to EMSAs. It should be noted that this is not an equilibrium technique: the current applied during gel

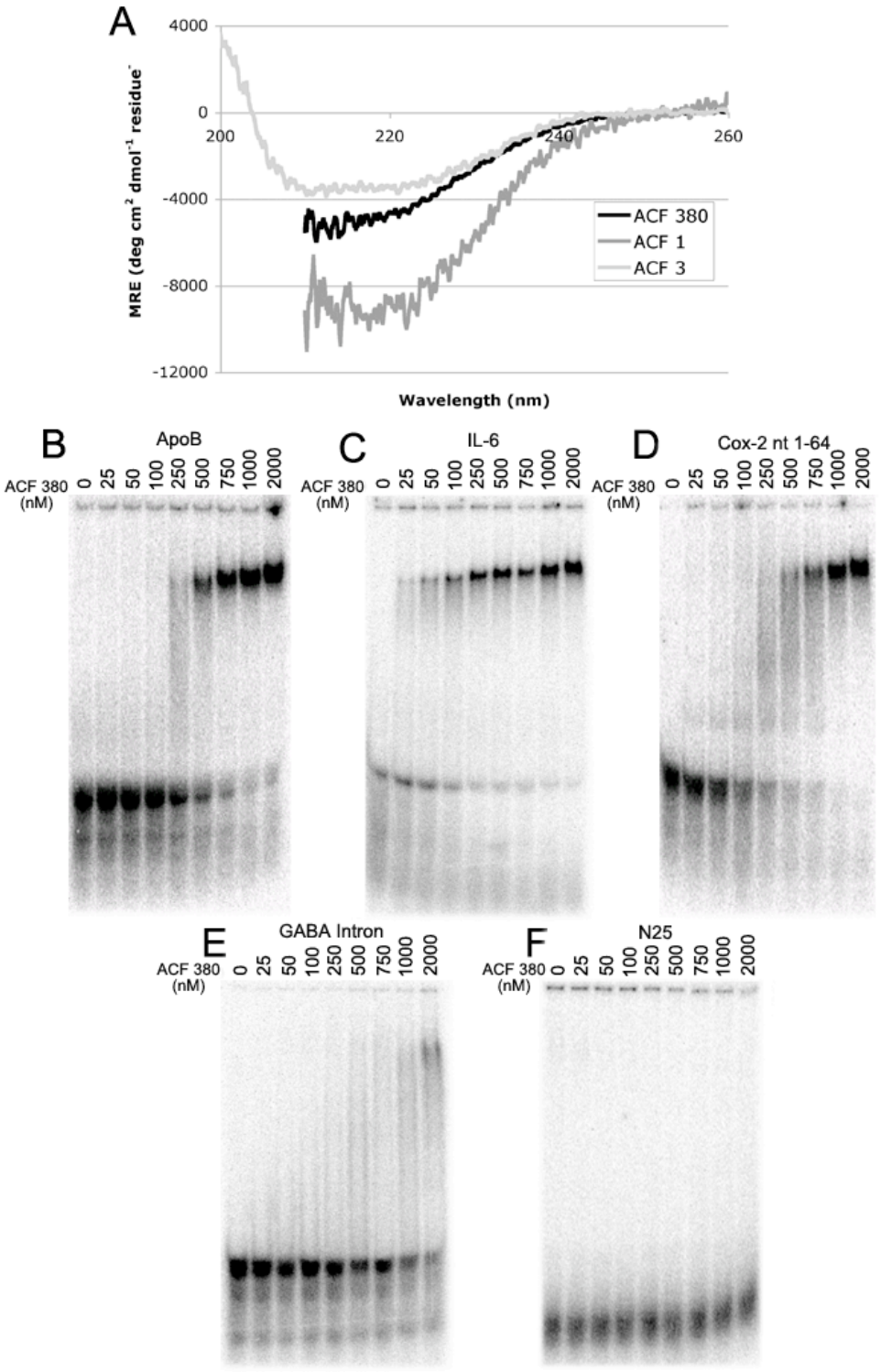


Figure 3.2 - ACF binding activity. A) UV CD spectra of ACF380, ACF1, and ACF3 plotted as mean residue ellipticity as a function of wavelength (nm). B-F) Purified protein in concentrations indicated were incubated in 100 mM KCl with *in vitro* transcribed, ^{32}P-RNA ApoB (B), IL-6 AU-rich (C), Cox-2 nt 1-64 (D), GABA Intron (E), N25 (F).

resolution can result in protein:RNA complex dissociation. For that reason, all gels were resolved at 4°C to stabilize complexes and reduce dissociation. With those caveats in mind, we felt EMSAs to be the best method to initially examine ACF interaction with a panel of five RNAs that explore the range of ACF's target specificity: ApoB, IL-6, Cox-2, GABA Intron, and N25 (Figure 3.1B). ApoB mRNA is ACF's canonical binding target [3, 9], while IL-6 and Cox-2 represent AU-rich RNAs that ACF stabilizes [10] and destabilizes, respectively. GABA Intron is an unstructured, CU-rich RNA [11] with no known or predicted physiological interaction with ACF; this target probes ACF's preference for AU-rich content. N25 is a random 25-nucleotide sequence pool to control for non-specific binding.

For all EMSA experiments, the binding buffer contained 100 mM KCl, which should approximate physiological salt concentrations and suppress nonspecific electrostatic interactions. The experiments showed that ACF380 binds ApoB mRNA: at 250 nM protein, the band corresponding to the complex is not well defined (it is smeared), but at 500 nM protein, there is a strong shifted band (Figure 3.2B). However, ACF380 appears to bind IL-6 and Cox-2 RNAs with higher affinity (Figure 3.2C-D), since complexes are first observed with only 25 nM protein. ACF380 binding to IL-6 RNA results in a single band with a mobility approximately equal to that of the ACF:ApoB complex. ACF380 binding to Cox-2 produces two complexes, perhaps corresponding to two ACF binding sites with different affinities. The lower band (poorly defined in the gel) is populated at lower protein concentrations and the upper shift appears at higher protein concentrations at or near the position of the single band in the ApoB and IL-6 gels. Although the stoichiometry has not been directly measured for any

complex, one interpretation of these data is that the first complex visible in the Cox-2 experiments represents a 1:1 protein:RNA stoichiometry, while the slow mobility complex in all experiments results from a 2:1 protein:RNA complex. Protein association to ApoB and IL-6 RNAs might be cooperative, or alternatively, the binding affinities of two sites could be equal, since there is no evidence of a first binding event for those complexes.

In contrast to ApoB, IL-6, and Cox-2 RNAs, the GABA Intron sequence is CU-rich, and its use here addresses whether pyrimidines alone are sufficient for ACF380 binding. As shown (Figure 3.2E), ACF does bind this RNA, but with far weaker affinity than to the AU-rich RNAs. Although a shifted band from the protein:RNA complex is not clearly visible in these experiments (except at 2 μM protein), there is significant depletion of the free RNA as ACF concentration is increased. There is also obvious smearing of the RNA with increasing ACF concentration, undoubtedly due to bound RNA being released from the complex during electrophoresis. Certainly ACF380 does bind GABA Intron RNA, although with weak affinity.

ACF380 failed to shift the N25 RNA pool at all concentrations examined (Figure 3.2F). This confirms that the RNA shifts described above are due to specific binding and not non-specific electrostatic interactions.

We have shown that ACF380 has the ability to bind a range of RNA sequences. This includes IL-6 and Cox-2 RNAs, both of which it binds with similar affinity, though more tightly than its canonical target, ApoB mRNA. We have also shown that ACF380 binds not only to AU-rich sequences or AUUUA pentamers, but also to C/U

polypyrimidine tracts. This wide array of ACF RNA targets leads us to conclude that there is no single high affinity consensus binding site to which ACF binds.

This broad range of targets is not unique to ACF; HuR also exhibits relatively high affinity for both U-rich and AU-rich sequences [12]. While extensive work has been done to identify ARE RNAs, such as the ARE-RNA database (ARED) [13-16], we suggest it may be useful to broaden the search for RNAs that are regulated at the post-transcriptional level by ARE-BPs to include a wider range of sequences that include U-rich RNAs.

RNA Binding of ACF's RRMs.

In order to identify which RRMs of ACF bind RNA we incubated ACF1 with the four RNA constructs bound by ACF380, and used EMSAs to assess binding activity. In 100 mM KCl, we found that ACF1 failed to form a complex with ApoB RNA, consistent with previous reports that ACF1 is not necessary for ApoB RNA interaction [3]. However, ACF1 does detectably bind to IL-6, Cox-2, and GABA Intron RNA (Figure 3.3A). For IL-6 RNA, two complexes were observed and appear to be equally populated (see arrows, Figure 3.3A). ACF1:Cox-2 forms a single fast mobility complex, while the GABA Intron complex has a slower mobility. These data suggest that ACF1 is involved in ACF's interaction with all three of these RNAs, but that its contribution to binding varies with the RNA target.

We also assayed ACF3 with the same RNA targets, but detected no complex formation (Figure 3.3B). This suggests that ACF3 is either not involved in RNA binding, or that its RNA target is currently undiscovered. It is not entirely surprising to find that

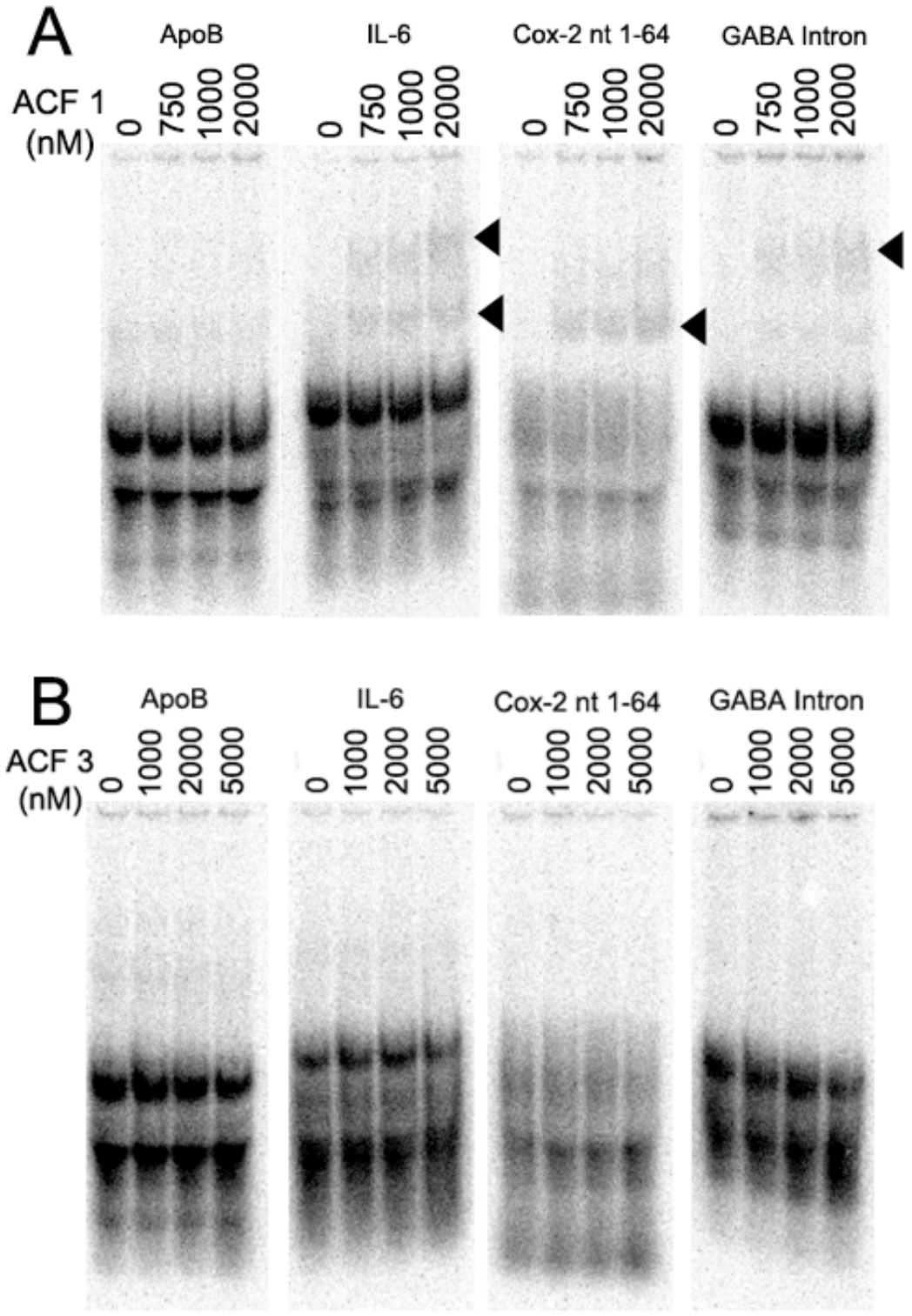


Figure 3.3 - EMSA of purified ACF1 (A) or ACF3 (B) in 100 mM KCl with *in vitro* transcribed, ^{32}P-RNAs as indicated.

this RRM fails to interact with RNA: this domain has a non-canonical RNP1, which may reduce protein:RNA points of contact. RRMs that fail to bind RNA are not unusual; the third RRM of HuD, which is highly homologous to ACF, does not bind RNA [17], nor does the second RRM of U1A [18]. It is possible that these RRMs bind RNA targets that have not yet been identified, but another possibility is that they participate in protein-protein interactions. We suggest that each ACF:RNA complex could involve a different arrangement of RNA:RRM interactions, as well as different associated proteins. It is possible that this combination of multiple RNA binding mechanisms, along with interactions with other ARE-BPs accounts for ACF's ability to differentially regulate various targets *in vivo*.

Although interactions with other proteins likely alter ACF binding outcomes in the cell, it is interesting that even in the absence of other proteins in our *in vitro* assays ACF appears to use different mechanisms to bind each of the RNAs examined, based on the apparent affinities and different complexes observed in EMSA experiments. ACF1 appears to be uninvolved in the binding of ApoB mRNA. Since ACF3 alone does not detectably bind to ApoB RNA, but ACF380 clearly does, either RRM2 alone or some synergistic coupling of the RRMs in the context of the entire protein must account for the observed binding. Conversely, ACF1 clearly participates in the binding of IL-6, Cox-2, and GABA Intron RNAs, but the apparent binding affinity of ACF1 alone is less than that of ACF380, again implicating RRM2 in association with all three RNAs.

ACF binds IL-6 RNA.

Due to the observation that each of the detected ACF:RNA interactions are unique, we took a closer look at ACF binding to IL-6 RNA. First, we used UV-crosslinking to examine the binding of full length ACF to the AU-rich region of the IL-6 3'UTR (Figure 3.4A). Recombinant full length ACF was incubated with ^{32}P-labeled IL-6 3'UTR AU-rich RNA in the presence of increasing concentrations of unlabeled IL-6 3'UTR AU-rich RNA. Protein-RNA complexes were then treated with RNaseT1 to degrade unprotected RNA, UV-crosslinked, and resolved by SDS-PAGE. The unlabeled RNA competed for binding to the labeled RNA at concentrations from 5x-10x molar excess (Figure 3.4B), consistent with ACF's high affinity for this construct (Figure 3.2C).

The AU-rich region of the IL-6 3'UTR contains 4 AUUUA pentamers that are possible ACF-binding sites. In order to more closely assess the ACF binding site on IL-6 RNA, we mutated these pentamers in pairs (Figure 3.4A). When full length ACF was UV-crosslinked to two of these mutants, it was obvious that mutation of the proximal pair of pentamers (nt 130-145) abrogates ACF binding (Figure 3.4C). However, mutation of the distal pair of pentamers (nt 179-196) has no obvious effect on ACF:IL-6 RNA interaction, with the caveat that these assays are not quantitative. To further investigate these interactions, we bound ACF to each of the radiolabeled mutant constructs in the presence of increasing molar excesses of unlabeled wild-type AU-rich RNA (Figures 3.4D and E). It is clear that what little ACF:Mutant 1 interaction was observed was competed by 1x concentration of wild-type AU-rich RNA. However, the ACF:Mutant 2 interaction was maintained until 5x-10x molar excess of wild-type AU-rich RNA, a range

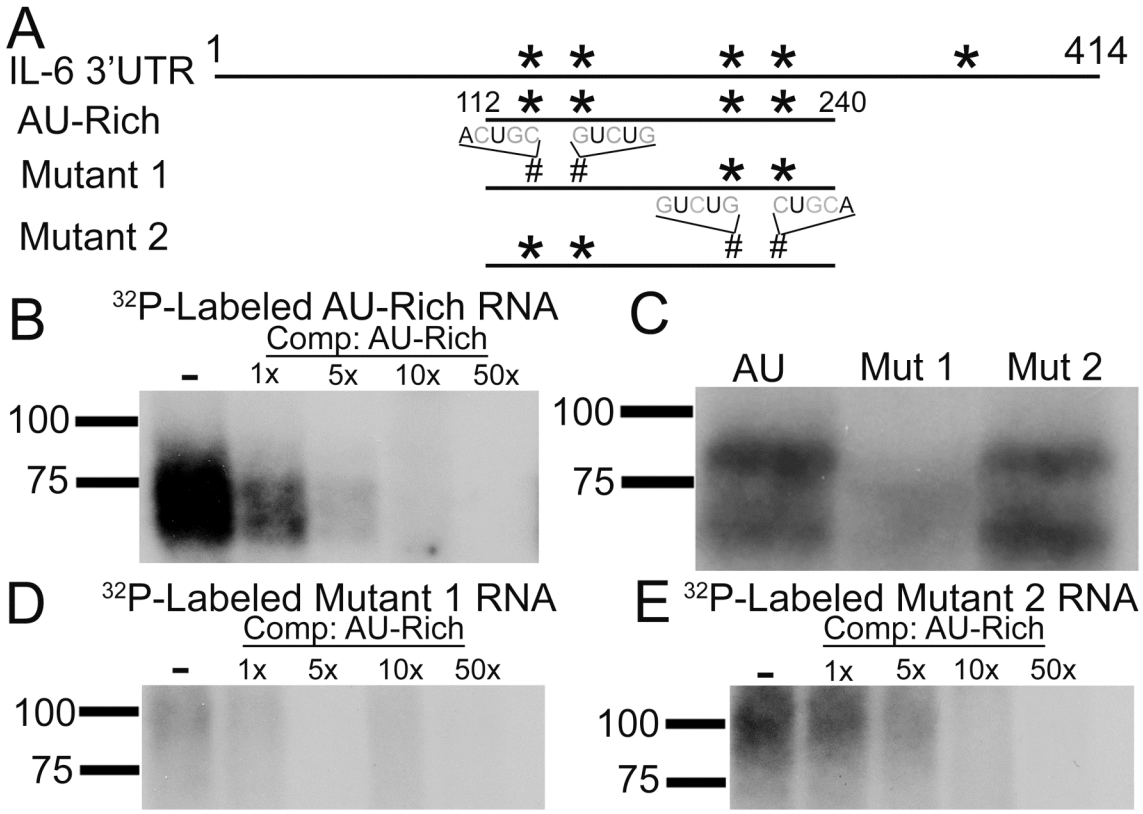


Figure 3.4 - ACF binding to IL-6 3'UTR RNA. A) Schematic of the IL-6 3'UTR. Numbering begins with first nucleotide of the 3'UTR. * indicates AUUUA pentamer. # indicates mutated pentamer. B) 40 nM purified full-length ACF was UV-crosslinked with *in vitro* transcribed ^{32}P-AU-rich RNA in the presence of 1x, 5x, 10x, or 50x molar excess of *in vitro* transcribed unlabeled AU-rich RNA at room temperature in 100 mM KCl with 3 mg/mL Heparin. Reactions were treated with 5 U/μL RNase T1, UV-crosslinked, resolved by 10% SDS-PAGE and exposed by autoradiography. C) ACF was UV-crosslinked to ^{32}P-AU-rich, Mutant 1, or Mutant 2 RNA under the same conditions as described above. D-E) 40 nM purified full-length ACF was UV-crosslinked with *in vitro* transcribed ^{32}P-labeled Mutant 1 (D) or Mutant 2 (E) RNA in the presence of 1x, 5x, 10x, or 50x molar excess of *in vitro* transcribed unlabeled AU-rich RNA under the same conditions described in (B).

similar to that of the AU-rich self-competition (Figure 3.4B), suggesting that ACF has similar affinity for both the wild-type and Mutant 2 RNAs.

These crosslinking data indicate that ACF specifically interacts with at least one of the pentamers in nt 130-145 of the IL-6 3'UTR. While we have not identified the exact ACF binding site, it is clear that ACF binds a discrete sequence within the 129 nt IL-6 3'UTR region that also confers ACF-dependent stabilization *in cellulo* (Chapter 2). While our EMSA data suggest ACF binds a wide range of targets and may not bind a specific consensus motif, it is clear that binding to this target RNA is isolated to a distinct site within the 3'UTR.

ACF Binding to Cox-2 RNA.

In order to more closely investigate the domains of ACF involved in Cox-2 RNA binding, we used a panel of recombinant ACF mutants to probe the ACF:Cox-2 interaction (Figure 3.1A). Six different purified recombinant ACF constructs were UV-crosslinked to ^{32}P-Cox-2 nt 1-64 RNA in buffer with 100 mM KCl (Figure 3.5). Wild-type ACF, ACF380, and ACFΔdsRBD all bound Cox-2 RNA. However, ACF constructs in which RRM1, RRM2, or RRMs 1-3 were deleted showed no interaction with the RNA. This is consistent with EMSA data suggesting that ACF1 is necessary but not sufficient to account for ACF:Cox-2 interaction (Figures 3.2B and 3.3A), and supports our conclusion that ACF RRM2 participates in RNA binding.

In an effort to elucidate the role of ACF RRM2 in Cox-2 RNA binding, we used its tryptophan fluorescence to probe for RNA binding in the context of ACF380. [Our attempts to express RRM2 alone in *E. coli* were not successful. The constructs RRM2,

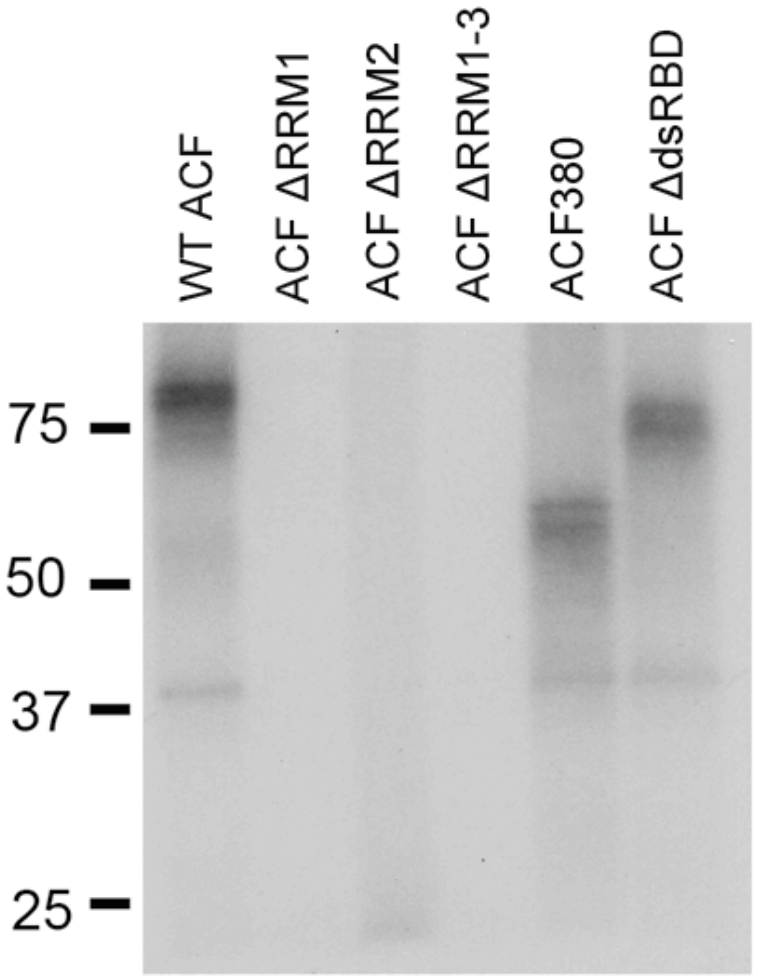


Figure 3.5 - *In vitro* transcribed, ^{32}P-labled Cox-2 nt 1-64 RNA was UV-crosslinked to 40 nM purified protein at room temperature in 100 mM KCl with 3 mg/mL Heparin. Reactions were treated with 5 U/μL RNase T1, UV-crosslinked, resolved by 10% SDS-PAGE and exposed by autoradiography.

RRM1-2, and RRM2-3 failed to express.] Tryptophan is a naturally fluorescent amino acid that is extremely sensitive to environmental changes. Therefore, the binding of an RNA molecule in close proximity (ie. base stacking) to a tryptophan can result in observable differences in fluorescence emission. ACF1 contains one naturally occurring tryptophan approximately 9 amino acids before the predicted start of the domain (trp 47). RRM2 contains two tryptophans, one in loop 5 between predicted α-helix 2 and β-strand 4 (trp 207), and the second at the end of predicted β-strand 4 (trp 215) (Figure 3.6A, structural predictions based on SWISS-MODEL homology model). By comparing the tryptophan fluorescence intensity of ACF1 and ACF380 in the presence and absence of RNA, we hoped to find evidence of RNA association with RRM1 and/or RRM2.

We observed the steady-state fluorescence emission spectrum of ACF380 (excited at 300 nm) to have a maximum at 337 nm. Fully solvent-accessible tryptophan has a fluorescence maximum of 348-350 nm; however, a tryptophan that is less solvent-accessible (ie. due to protein packing) is expected to have a maximum emission at a shorter wavelength. Keeping in mind that this spectrum has contributions from all three tryptophan residues, upon the addition of equimolar Cox-2 nt 1-64 RNA, the fluorescence intensity of the protein decreased by ~17%, with no shift in the emission maximum (Figure 3.6B). This response clearly shows that RNA binding quenches the fluorescence of one or more of the tryptophans in ACF380.

In order to identify which of ACF380's tryptophans are quenched upon RNA binding, we repeated the experiment using ACF1, which has one tryptophan in common with ACF380 (trp 47). The steady-state fluorescence emission spectrum of ACF1 alone (excited at 300 nm) has a maximum at 339 nm, comparable to that of ACF380; neither

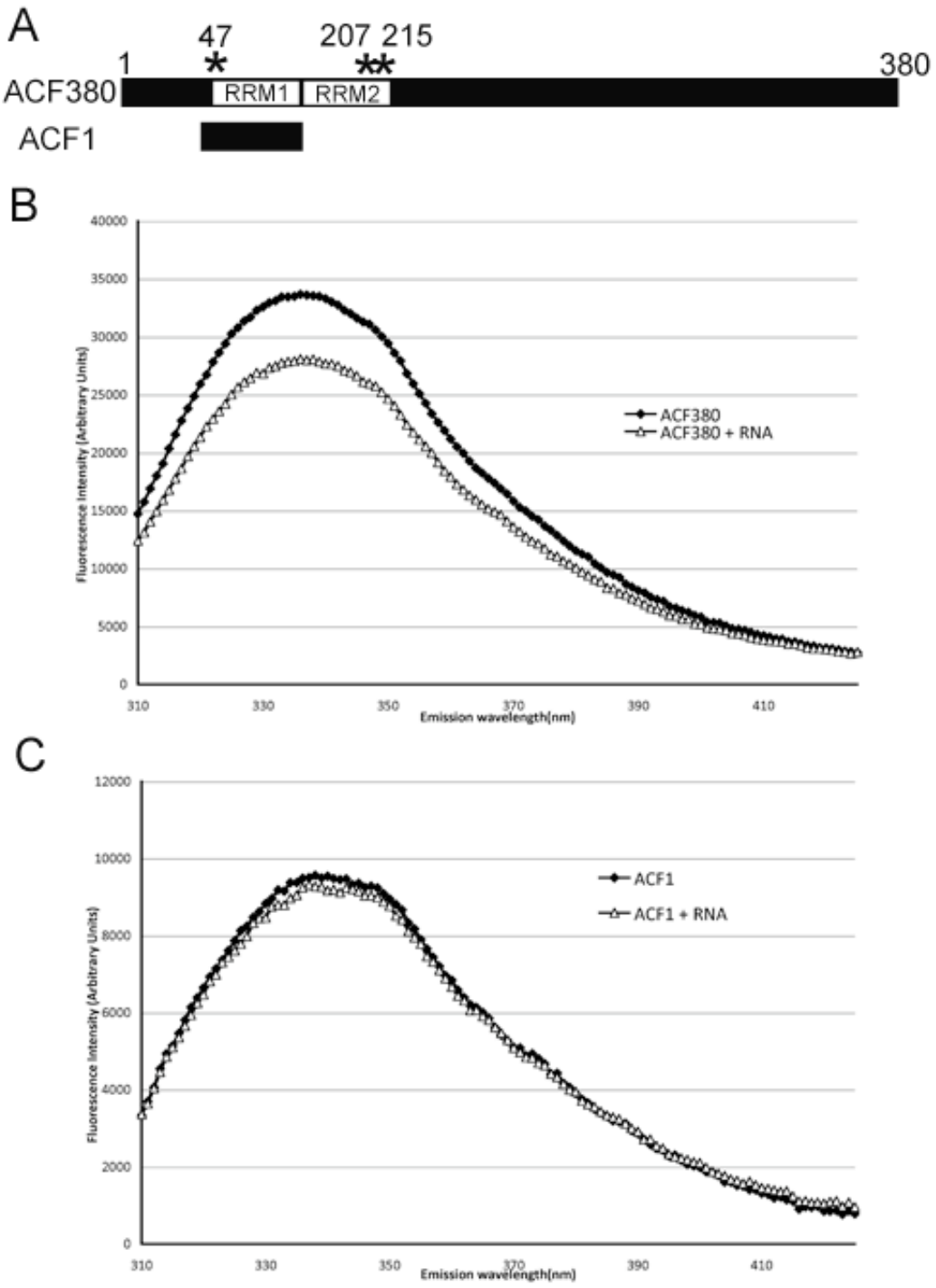


Figure 3.6 - ACF Fluorescence. A) Schematic of ACF380 and ACF1 with tryptophans (*) indicated. Emission spectra for 1μM ACF380 (B) or ACF1 (C) in the presence or absence of *in vitro* transcribed Cox-2 nt 1-64 RNA of equal concentration (1 μM) in 100 mM KCl. Excitation was at 300 nm.

the emission maximum nor the emission intensity changed in the presence of equimolar Cox-2 nt 1-64 RNA (Figure 3.6C). This response indicates that trp 47 does not contribute to the fluorescence quenching of ACF380 upon RNA binding. Therefore, RNA binding must result in significant quenching of tryptophans 207 and/or 215 in RRM2, strongly suggesting that this domain of ACF interacts with Cox-2 RNA.

These binding data indicate that while ACF1 is necessary for ACF:Cox-2 interaction (Figure 3.5), it is not sufficient; the same is also true for RRM2, which is implicated in Cox-2 binding through tryptophan quenching. It is not unique for two RRMs to participate in RNA interaction [11]. However, this case is of particular interest as both our EMSA data (Figure 3.3A) and previous work showing that ACF ΔRRM1 can UV-crosslink to ^{32}P-ApoB mRNA [3] suggest RRM1 does not participate in the ACF:ApoB interaction. This clearly reveals that ACF binds its RNA targets by more than one mechanism. We propose this could provide a means for ACF's ability to interact with multiple targets *in vivo* and differentially regulate their stability. It is also possible that the varied mechanisms of ACF:RNA interaction allow or define ACF's ability to interact with other proteins, which, once again, may contribute to the differential effect ACF has on the stability of its targets.

In an effort to more closely examine the ACF:Cox-2 interaction, we undertook mutation of Cox-2 3'UTR nt 1-64 to determine if ACF exhibits a preference for a particular binding site, as observed in the IL-6 3'UTR. We introduced U to C mutations in three separate AU-rich (ARE) regions of the construct (Figure 3.7A). When full-length recombinant ACF was UV-crosslinked to wild-type or mutant Cox-2 RNAs, it was clear that ACF's primary binding site on this construct is the "B" region (nt 37-46). Keeping in

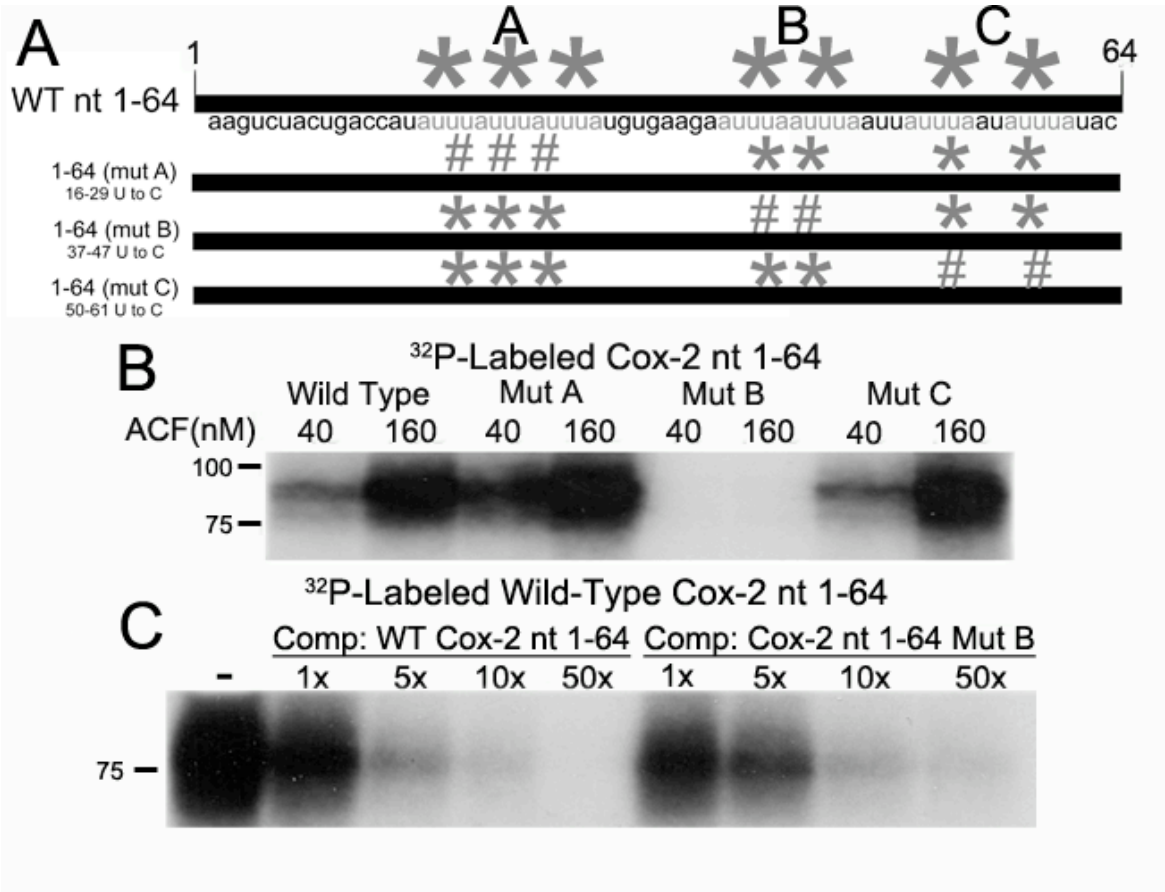


Figure 3.7 - ACF exhibits a preferred binding site in Cox-2 nt 1-64. A) Schematic of mutations in Cox-2 nt 1-64. * indicates AUUUA sequence, # indicates U to C mutated sequence. B) Purified full-length ACF (concentration as noted) was UV-crosslinked with *in vitro* transcribed ^{32}P-RNA at room temperature in 100 mM KCl with 3 mg/mL Heparin. Reactions were treated with 5 U/μL RNase T1, UV-crosslinked, resolved by 10% SDS-PAGE and exposed by autoradiography. C) 40 nM ACF was UV-crosslinked to ^{32}P-Cox-2 nt 1-64 (WT) in the presence of 1x, 5x, 10x, or 50x molar excess of *in vitro* transcribed unlabeled WT or Mut B Cox-2 nt 1-64 RNA under the same conditions as in (B).

mind that these assays are not quantitative, it appears mutation of the A and C regions had little effect on ACF interaction while mutation of the B region entirely abrogated crosslinking (Figure 3.7B). To further assess ACF's affinity for this region, we used cold-competition assays to compare ACF binding of wild-type and Mutant B Cox-2 nt 1-64 RNAs. ACF was incubated with ^{32}P-wild-type Cox-2 nt 1-64 RNA in the presence of increasing molar excess of unlabeled wild-type or Mutant B Cox-2 nt 1-64 RNA. While unlabeled wild-type RNA is able to compete for bound ^{32}P-Cox-2 between 1x and 5x molar excess, unlabeled Mutant B RNA requires at least 5-10x molar excess (Figure 3.7C). This indicates two important points: ACF has a higher affinity for wild-type RNA than for Mutant B, and while ACF prefers to bind region B it is also capable of interacting with regions A and C. It is important to note that the B region is consistent with the mutation used in Cox-2 half-life analysis that confers mRNA destabilization (Chapter 2). This suggests that ACF interaction with region B is at least partly responsible for ACF-dependent destabilization of Cox-2 mRNA *in cellulo*.

Due to ACF's wide range of sequence preference and possible multiple binding sites on Cox-2 3'UTR nt 1-64, we examined the need for structural specificity. A fragment of the Cox-2 3'UTR was generated consisting of nt 37-64, but with all uridines in region C changed to cytosines (Figure 3.8A), leaving only one ACF binding site (region B). This RNA was then manipulated by further 3' or 5' mutations to create different structural contexts for region B: (a) ssRNA, (b) dsRNA, or (c) a hairpin loop (Figure 3.8B), using *mfold* [1, 2] to predict the energetically most favorable structure. Using UV-crosslinking with purified recombinant full-length ACF, we determined that ACF binds the ssRNA construct, while interaction is completely abrogated when the

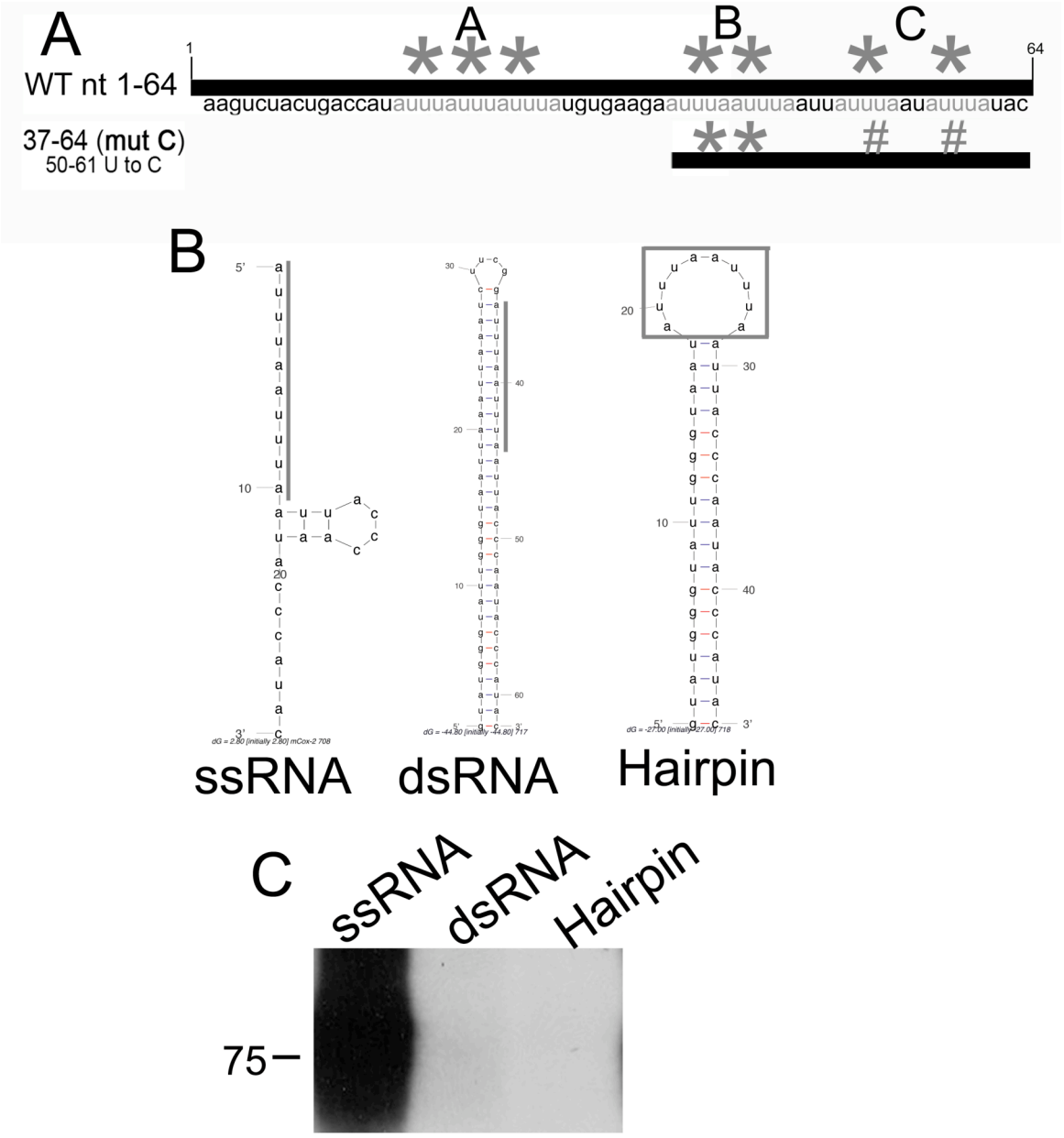


Figure 3.8 - ACF exhibits a preference for ssRNA. A) Schematic of Cox-2 nt 37-64 with U to C mutations in nt 48-64. * indicates AUUUA sequences, # indicates U to C mutation. B) The mutant sequence described above was generated in the context of three predicted secondary structures (*mfold*, [1, 2]): ssRNA, dsRNA, hairpin loop, with ACF binding site "B" indicated. C) 40 nM purified recombinant ACF was UV-crosslinked to *in vitro* transcribed ^{32}P-RNA at room temperature in 100 mM KCl with 3 mg/mL Heparin. Reactions were treated with 5 U/μL RNase T1, UV-crosslinked, resolved by 10% SDS-PAGE and exposed by autoradiography.

binding site is double-stranded or in a hairpin loop (Figure 3.8C). It should be noted that we tried only one hairpin loop structure with a ten nucleotide loop. It is feasible that ACF may bind stem-loop RNA of different sizes. Still, these data indicate that not only the RNA sequence but also its structure is important for ACF binding. This structural preference may be the basis for ACF's ability to differentiate between RNAs or between sites on a single RNA, such as binding Cox-2 region B instead of regions A or C.

Though our data suggest that ACF prefers a ssRNA binding site on Cox-2 RNA, it's binding site on ApoB mRNA in the editing complex is believed to form a stem-loop structure (Figure 3.9A) [19, 20]. When examined in the context of Cox-2 3'UTR nt 1-64, multiple structures proposed by *mfold* situate "region B" in a ssRNA context (Figure 3.9B). Given that ACF380 exhibits higher affinity for Cox-2 3'UTR nt 1-64 than ApoB RNA (Figure 3.2B and D), we predict that the ApoB secondary structure (Figure 3.9A) is not ACF's preferred structure (though we acknowledge we have used ACF380 in these assays and full length ACF may exhibit altered affinity for one or both targets). However despite this sub-optimal binding site, ACF is known to interact with Apobec-1 (which also binds ApoB mRNA) at the editing site, and it is possible that this interaction increases ACF's affinity for ApoB RNA. This could due to cooperative binding of both proteins to the RNA or a change in ACF's affinity for the RNA upon Apobec-1 interaction. For example, U2B" is an RRM-protein which binds a single biological target (U2 snRNA stem-loop IV) in complex with U2A' (which does not detectably bind RNA) [21]. Alone, U2B" binds its target RNA very weakly, but the U2B":U2A' complex results in a highly specific interaction with U2 snRNA stem-loop IV. It is feasible to

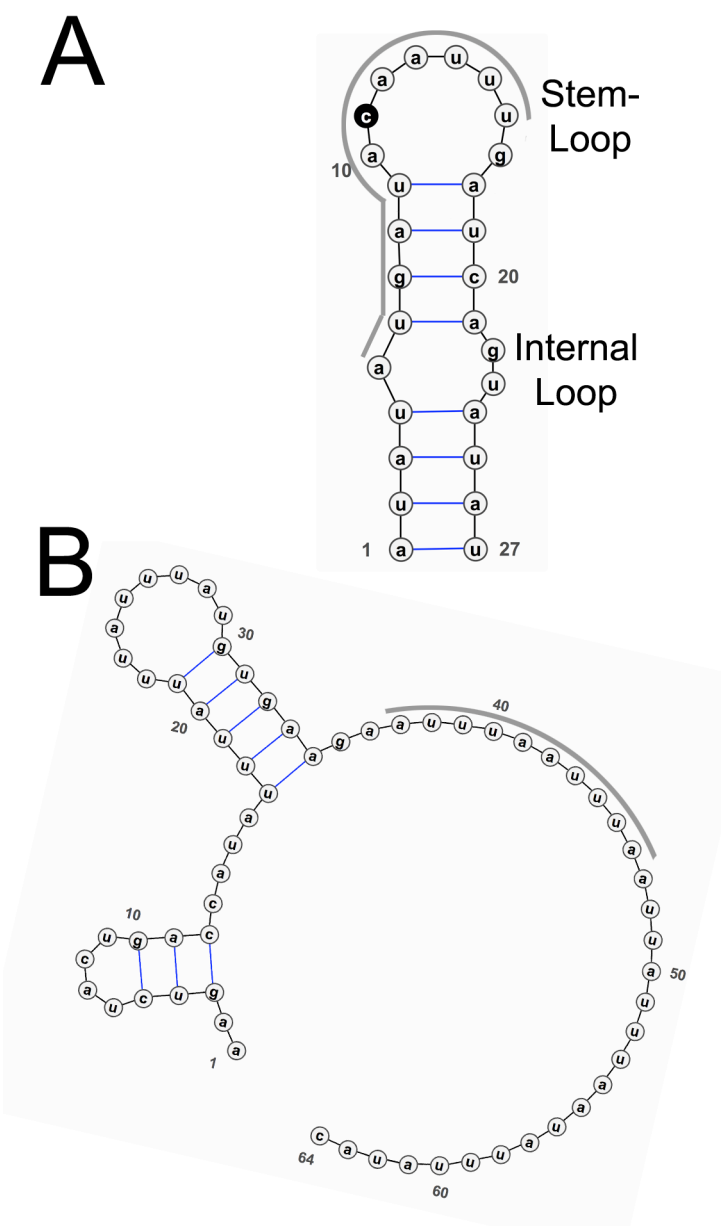


Figure 3.9 - Predicted secondary structures of ACF targets. Regions marked indicate ACF binding sites. A) Structure of minimal ApoB mRNA editing cassette with edited C highlighted. B) Representative structure of Cox-2 3'UTR nt 1-64 as predicted by *mfold.*

consider that the ACF:Apobec-1 interaction may enhance the ACF:ApoB RNA interaction.

Alternatively, published work has suggested that upon loosely binding to ApoB mRNA, ACF may result in melting of the ApoB mRNA stem-loop structure [19]. While there are currently few data to support this hypothesis, it is clear that the stem between the internal loop and stem-loop (Figure 3.9A) has a low melting temperature as it is comprised of mainly A-U pairs. ACF may transiently interact with un-paired nucleotides in the loops but more fully bind the ssRNA upon inducing the stem to melt. This is not a unique characteristic of an ARE-BP; Auf1 has also been observed to alter target RNA secondary structure upon binding [22]. However, it is important to remember that ACF appears to interact with each of the RNAs assayed differently; therefore more work is required to ascertain the method by which ACF contributes to both editing and the stability of its other targets.

Fluorescence and UV-crosslinking data indicate that, in the case of Cox-2, ACF exhibits a distinct preference for a specific, ssRNA binding site (nt 37-46), which appears to confer a robust phenotype in the destabilization of reporter RNA (Chapter 2). It is reasonable to hypothesize that this specific ACF:Cox-2 interaction may be further facilitated by interaction with other ARE-BPs; ACF may bind the ssRNA target on Cox-2 and proceed to recruit other ARE-BPs. It has been well established that ARE-RNAs are bound by multiple ARE-BPs, often with additive or competing effects [23-25]. This is very likely to be true in the case of ACF which is known to bind at least one ARE-BP, Apobec-1 [9], and we have shown to interact with both HuR and Auf1 (Chapter 2).

Conclusion.

RRM proteins exhibit a wide range of binding behaviors. Some, such as U1A, bind single sequence-specific targets [26], while others, such as HuR, bind a large number of targets with similar sequence characteristics [12]. Still others, such as U2B" bind sequence-specific targets, but require the assistance of a co-factor for high-affinity binding [21]. Therefore, it was necessary to undertake biochemical assessment of ACF's binding activity to understand how this RRM protein interacts with its target RNAs. Our data indicate that ACF binds a wide range of AU and U-rich RNAs with variable affinity, but exhibits a preference for ssRNA. We have also shown that while ACF does not appear to bind a single consensus motif, it binds each target at a discrete site that may confer ACF's ability to regulate stability. Finally, we have assessed the role of ACF's individual RRMs in RNA-binding. ACF binds each of the RNAs examined in a unique manner that requires some combination of RRMs 1 and 2, indicating that these two RRMs may fold in close proximity or synergistically. We propose that the physiological impact of ACF's RNA binding activity *in vivo* is regulated both by the binding behaviors we have observed as well as interaction with other proteins to result in a net mRNA stabilization or destabilization phenotype.

MATERIALS AND METHODS

Cloning and Expression of Recombinant proteins – Full length ACF, ACF380, ACF ΔRRM1, ACF ΔRRM2, ACF ΔRRMs 1-3, and ACF ΔdsRBD were all cloned and expressed as self-cleaving intein fusion proteins as previously described [3, 10].

6xHis-ACF1 protein was expressed in BL-21(DE3) (Invitrogen) cells for 4 hours at 37°C after 1mM IPTG induction. Cells were pelleted and frozen at -70°C overnight. Cell pellet was reconstituted in 20 mM sodium acetate pH 5.3, 50 mM NaCl, 2 mM EDTA, 600 U DNAseII, 0.8 mg/mL PMSF and lysed by French press. Lysate was centrifuged and dialyzed against 10 mM Imidazole Buffer (50 mM NaH_2PO_4 pH 8.0, 500 mM NaCl, 10 mM Imidazole, 10 mM β-mercaptoethanol (BME), 0.05% Triton-X). Dialysate was incubated with 6 mL Ni-NTA beads (Qiagen, prepared as recommended by manufacturer) at 4°C for 1 hour with shaking. Beads were subsequently washed 3 times with 5 volumes of 10 mM Imidazole Buffer and 2 times with 5 volumes of 50 mM Imidazole Buffer. 6xHis-ACF1 was eluted from the beads in 20 mL 300 mM Imidazole buffer. Elutate was dialyzed against 50 mM NaH_2PO_4, 200 mM NaCl, 10 mM BME, 5 mM sodium citrate. Dialysate was incubated with 250 U Pro-TEV protease, which contains a His-tag (Promega), at room temperature for 14 hours. Cleaved protein was incubated with 5 mL Ni-NTA beads at 4°C for 20 minutes to remove the protease and cleaved tag, and the supernatant retained. The supernatant was again exposed to 5 mL Ni-NTA beads at 4°C for 1 hour and the supernatant retained. Supernatant was concentrated in a 3000 MWCO Vivaspin and subsequently dialyzed against 20 mM HEPES pH 8.0, 100 mM KCl, 20 mM EDTA, 20% Glycerol, 2 mM TCEP. Final protein concentration was determined by Bio-Rad Protein Assay.

ACF3 was PCR amplified and subcloned into pTAC2 using the NcoI/BglII sites. Protein was expressed in BL-21(DE3) cells for 3 hours at 37°C following 1mM IPTG induction. Cell pellet was reconstituted in 20 mM sodium acetate pH 5.3, 50 mM NaCl, 2 mM EDTA, 600 U DNAseII, 0.8 mg/mL PMSF and lysed by French press. Lysate was

centrifuged and dialyzed against 50 mM Tris pH 7.5, 10 mM NaCl, 0.2 mg/mL PMSF. Dialysate was then loaded on a Bio-Rad Q Column pre-equilibrated with 50 mM Tris pH 7.5 and eluted from the column by a gradient containing 10-500 mM NaCl. Fractions were concentrated in 3000 MWCO Vivaspin and the concentration determined spectrophotometrically. For ACF3, $\varepsilon_{280} = 2560\ M^{-1}cm^{-1}$.

Circular Dichroism Spectra – CD spectra were recorded with a Jasco J715 instrument at room temperature. ACF380 and ACF1 experiments were conducted in 20 mM Hepes pH 8.0, 100 mM KCl, 2 mM EDTA, 2 mM TCEP, 20% Glycerol. ACF3 experiments were conducted in 50 mM Tris pH 7.5, 100 mM NaCl. Proteins were assayed at the following concentrations: ACF380 = 4.74 μM, ACF1 = 3 μM, ACF3 = 22.5 μM.

Cloning of RNAs – The wild-type Cox-2 nt 1-64 construct in pcDNA3 was a generous gift from Dr. Aubrey Morrison (Washington University in St. Louis). For all other Cox-2 constructs, top and bottom oligonucleotides were annealed with the appropriate cohesive ends and ligated into pcDNA3 in the KpnI/BamHI sites. Cloning of ApoB [27], IL-6 3'UTR AU-Rich [10], and GABA Intron γ57 [11] plasmids has been previously described.

In vitro RNA synthesis – RNA constructs for EMSA and gel shift were enzymatically synthesized by T7 RNA polymerase. ApoB plasmid was linearized by HindIII digestion. IL-6 and Cox-2 plasmids were linearized by BamHI digestion. GABA Intron plasmid was linearized by AccI digestion. N25 was transcribed directly from a 25 nt random pool template. ApoB, IL-6 and Cox-2 RNAs were internally labeled with [α^{32}P]UTP and GABA Intron and N25 RNAs were internally labeled with [α^{32}P]UTP and [α^{32}P]CTP. All RNAs were gel purified prior to use. For fluorescence binding experiments, wild type

Cox-2 nt 1-64 RNA was transcribed, gel purified, and quantified by absorbance at 260 nm.

Electrophoretic Mobility Shift Assays – [α^{32}P]-labeled RNA was heated to 95°C for 5 minutes, then supplemented with 5 mM KCl and 10 μg/mL yeast tRNA and quenched on ice. Reactions were carried out in 10 mM Hepes pH 8.0, 100 mM KCl, 0.1 mM EDTA, 10 mM DTT, 20 μg/mL BSA, 0.1μL RNAsin per reaction (Promega). The RNA:protein samples were incubated at room temperature for 15 minutes, then glycerol was added and the total sample loaded on 4% polyacrylamide gels (29:1 acrylamide:bisacrylamide ratio in 100 mM Tris/100 mM Glycine, pH 8.3). Gels were run at 4°C at 8 V/cm for 2-3 hours. Data were analyzed via phosphoimager with ImageQuant software.

UV-Crosslinking Assays – [α^{32}P]-labeled RNA was heated to 70°C for 5 minutes then supplemented with 5 mM KCl and quenched on ice. Reactions were carried out in 10 mM Hepes pH 8, 100 mM KCl, 0.1 mM EDTA, 0.25 mM DTT, 2.5% glycerol and incubated at room temperature for 15 minutes. Heparin was added to 3 mg/mL and incubated at room temperature for 5 minutes. RNase T1 was added to 5 U/μL and incubated at room temperature for 5 min. Reactions were UV-irradiated on ice in a Stratalinker (Stratagene) at 250 millijoules/cm^2 then analyzed by 10% SDS-PAGE and visualized by autoradiography.

Fluorescence measurements – The steady state fluorescence emission spectrum for each protein was measured with a Photon Technology International 810 instrument. Tryptophans were excited at 300 nm and non-polarized emission intensity scanned from 310 to 425 nm with integration of 2 seconds. ACF380 data represents an average of 2 scans and ACF1 data represents an average of 4 scans. All samples were taken at room

temperature in 20 mM Hepes pH 8.0, 100 mM KCl, 2 mM EDTA, 2 mM TCEP, 20% Glycerol in BSA-blocked cuvettes.

REFERENCES

1 M. Zuker. (2003) Mfold web server for nucleic acid folding and hybridization prediction, Nucleic Acids Res. **31**, 3406-3415

2 D. H. Mathews, M. D. Disney, J. L. Childs, S. J. Schroeder, M. Zuker and D. H. Turner. (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure, Proc Natl Acad Sci U S A. **101**, 7287-7292

3 V. Blanc, J. O. Henderson, S. Kennedy and N. O. Davidson. (2001) Mutagenesis of apobec-1 complementation factor reveals distinct domains that modulate RNA binding, protein-protein interaction with apobec-1, and complementation of C to U RNA-editing activity, J Biol Chem. **276**, 46386-46393

4 M. P. Sowden, D. M. Lehmann, X. Lin, C. O. Smith and H. C. Smith. (2004) Identification of novel alternative splice variants of APOBEC-1 complementation factor with different capacities to support apolipoprotein B mRNA editing, J Biol Chem. **279**, 197-206

5 K. Arnold, L. Bordoli, J. Kopp and T. Schwede. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling, Bioinformatics. **22**, 195-201

6 L. Bordoli, F. Kiefer, K. Arnold, P. Benkert, J. Battey and T. Schwede. (2009) Protein structure homology modeling using SWISS-MODEL workspace, Nat Protoc. **4**, 1-13

7 T. Nagata, Y. Muto, M. Inoue, T. Kigawa, T. Terada, M. Shirouzu and S. Yokoyama. (2005) Solution structure of the RNA recognition motif of human APOBEC-1 complementation factor, ACF, Protein Database, pdb.org.

8 C. A. Galloway, A. Kumar, J. Krucinska and H. C. Smith. (2010) APOBEC-1 complementation factor (ACF) forms RNA-dependent multimers, Biochem Biophys Res Commun. **398**, 38-43

9 A. Mehta, M. T. Kinter, N. E. Sherman and D. M. Driscoll. (2000) Molecular cloning of apobec-1 complementation factor, a novel RNA-binding protein involved in the editing of apolipoprotein B mRNA, Mol Cell Biol. **20**, 1846-1854

10 V. Blanc, K. J. Sessa, S. Kennedy, J. Luo and N. O. Davidson. (2010) Apobec-1 complementation factor modulates liver regeneration by post-transcriptional regulation of interleukin-6 mRNA stability, J Biol Chem. **285**, 19184-19192

11 C. Clerte and K. B. Hall. (2006) Characterization of multimeric complexes formed by the human PTB1 protein on RNA, RNA. **12**, 457-475

12 I. Lopez de Silanes, M. Zhan, A. Lal, X. Yang and M. Gorospe. (2004) Identification of a target RNA motif for RNA-binding protein HuR, Proc Natl Acad Sci U S A. **101**, 2987-2992

13 T. Bakheet, M. Frevel, B. R. Williams, W. Greer and K. S. Khabar. (2001) ARED: human AU-rich element-containing mRNA database reveals an unexpectedly diverse functional repertoire of encoded proteins, Nucleic Acids Res. **29**, 246-254

14 T. Bakheet, B. R. Williams and K. S. Khabar. (2003) ARED 2.0: an update of AU-rich element mRNA database, Nucleic Acids Res. **31**, 421-423

15 T. Bakheet, B. R. Williams and K. S. Khabar. (2006) ARED 3.0: the large and diverse AU-rich transcriptome, Nucleic Acids Res. **34**, D111-114
16 A. S. Halees, R. El-Badrawi and K. S. Khabar. (2008) ARED Organism: expansion of ARED reveals AU-rich element cluster variations between human and mouse, Nucleic Acids Res. **36**, D137-140
17 X. Wang and T. M. Tanaka Hall. (2001) Structural basis for recognition of AU-rich element RNA by the HuD protein, Nat Struct Biol. **8**, 141-145
18 J. Lu and K. B. Hall. (1995) An RBD that does not bind RNA: NMR secondary structure determination and biochemical properties of the C-terminal RNA binding domain from the human U1A protein, J Mol Biol. **247**, 739-752
19 C. Maris, J. Masse, A. Chester, N. Navaratnam and F. H. Allain. (2005) NMR structure of the apoB mRNA stem-loop and its interaction with the C to U editing APOBEC1 complementary factor, RNA. **11**, 173-186
20 N. Richardson, N. Navaratnam and J. Scott. (1998) Secondary structure for the apolipoprotein B mRNA editing site. Au-binding proteins interact with a stem loop, J Biol Chem. **273**, 31707-31717
21 S. R. Price, P. R. Evans and K. Nagai. (1998) Crystal structure of the spliceosomal U2B"-U2A' protein complex bound to a fragment of U2 small nuclear RNA, Nature. **394**, 645-650
22 G. M. Wilson, K. Sutphen, M. Moutafis, S. Sinha and G. Brewer. (2001) Structural remodeling of an A + U-rich RNA element by cation or AUF1 binding, J Biol Chem. **276**, 38400-38409
23 N. Chang, J. Yi, et al. (2010) HuR uses AUF1 as a cofactor to promote p16INK4 mRNA decay, Mol Cell Biol. **30**, 3875-3886
24 B. C. Blaxall, A. Pende, S. C. Wu and J. D. Port. (2002) Correlation between intrinsic mRNA stability and the affinity of AUF1 (hnRNP D) and HuR for A+U-rich mRNAs, Mol Cell Biochem. **232**, 1-11
25 A. Lal, K. Mazan-Mamczarz, T. Kawai, X. Yang, J. L. Martindale and M. Gorospe. (2004) Concurrent versus individual binding of HuR and AUF1 to common labile target mRNAs, Embo J. **23**, 3092-3102
26 W. T. Stump and K. B. Hall. (1995) Crosslinking of an iodo-uridine-RNA hairpin to a single site on the human U1A N-terminal RNA binding domain, RNA. **1**, 55-63
27 F. Giannoni, D. K. Bonen, T. Funahashi, C. Hadjiagapiou, C. F. Burant and N. O. Davidson. (1994) Complementation of apolipoprotein B mRNA editing by human liver accompanied by secretion of apolipoprotein B48, J Biol Chem. **269**, 5932-5936

CHAPTER IV.

ACF Structural Predictions Using Homology Modeling.

Abstract.

Very little structural analysis of ACF has been undertaken. Unpublished data from a high-throughput group (RIKEN) has provided the structure of ACF RRM 3 by solution NMR (Figure 4.1). However, our data in the previous chapters suggest that RRM 3 plays no role in the interaction of ACF with any of the RNAs we examined. This leaves us with no experimentally determined structural data for ACF RRMs 1 and 2, which comprise ACF's RNA binding domains. Structural analysis can be highly informative in the study of a protein's function. Particularly in the case of ACF, we know that RRMs 1 and 2 work together in some fashion to interact with the RNAs in question. Without an idea of the secondary and tertiary structure of ACF it will be difficult to pursue more detailed analysis of any ACF:RNA interactions. For this reason, we turned to homology modeling via SWISS-MODEL Workspace [1, 2] to investigate possible structures ACF may adopt.

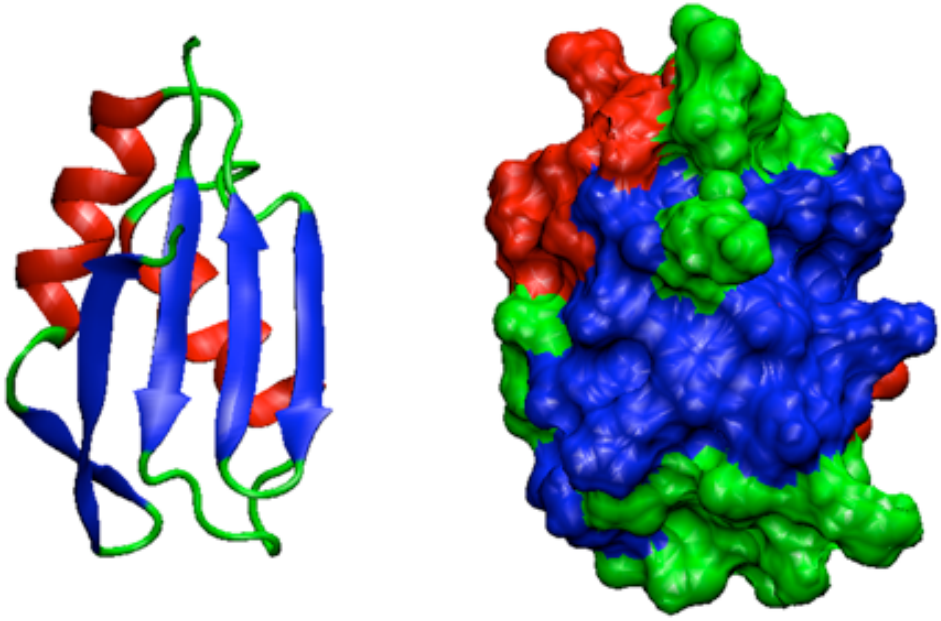


Figure 4.1 - Solution structure of ACF RRM3 in ribbon (left) and space-filling (right) representations. Structure as per RIKEN group, Nagata et. al (pdb: 2cpd) [3]. Colors represent secondary structure elements: red is alpha-helix, blue is beta-strand, green is turns and loops.

Homology Modeling.

SWISS-MODEL is one of a collection of web servers that generate homology models for proteins that have not been studied experimentally based on the structures of homologous proteins for which experimental structural data exist. The Protein Database (pdb.org) online repository for structural data currently contains just under 67,500 protein entries. However, UniProt KB (uniprot.org) the Protein Knowledgebase for protein sequence and functional information contains over 6 million entries. Clearly there is quite a backlog of proteins for which structural data needs to be experimentally determined. However, the information gained from currently known structures can be extrapolated to predict possible structural attributes of proteins that share sequence homology. Using homology modeling, otherwise known as comparative protein structure modeling, *in silico* models of a protein's 3-dimensional structure can be generated based on sequence alignment with a protein of known structure.

Homology modeling is based on four steps: 1) identification of proteins of known structure (template) that are evolutionarily related to the protein in question (target), 2) mapping of conserved and corresponding residues based on sequence alignment, 3) building a three-dimensional model, 4) model quality evaluation. Each of these steps is vitally important in order to achieve the most accurate model possible. However, it is important to acknowledge that even the best results are only models and cannot be relied on too heavily for interpretation of data.

While target proteins may be equally homologous to a number of potential template proteins, the three-dimensional model is generated using a single template. For that reason template selection is a very important step in the modeling process, as models

built from different templates can vary significantly. Template selection is based on three primary criteria: A) level of sequence similarity, B) quality of the solved structure, C) presence of ligands/co-factors in the solved structure [2]. The reliability and quality of an *in silico* model is contingent on the evolutionary distance between the target and the template. There is a direct correlation between the identity of the two aligned sequences and structural/positional deviation of the Cα atoms of their cores. As a rule, 50% identity between the template and target correlates to approximately 1.0 Å root mean square deviation of template Cα atoms from the target [2]. However, sequence similarity is not the only factor in selecting a template. Template quality can drastically affect the model outcome. The SWISS-MODEL Repository [4, 5] is a collection of potential templates extracted from the PDB that has eliminated unreliable entries such as theoretical models or low quality structures (i.e. those providing only Cα positions). From within this collection, identification of a specific target further takes into account potential quaternary structure of both the template and the target as well as other quality indicators such as empirical force field energy and ANOLEA mean force potential scores [6, 7].

Once a template is selected, the next step is to align the two sequences as accurately as possible. Alignment is a key step in ensuring a realistic three-dimensional model. This is even more important for low-identity sequences. Target/template pairs with less than 30% sequence identity are considered to be in the "twilight zone." These matches can still produce reliable or feasible models, but alignment errors and incorrect modeling of large insertions and/or deletions can cause significant inaccuracies in the model. By default SWISS-MODEL uses local pair-wise alignment [6], but users are highly encouraged to assess the alignment and make modifications as necessary. It is also

advisable to use other alignment algorithms (ie. ClustalW2, [8]) to look for differences in the alignments. Any alignment can be uploaded to SWISS-MODEL for use in modeling; it can be advantageous to examine the variations in models generated from multiple alignments of the same two sequences in order to assess the differences in the secondary and tertiary structures.

Once a template is selected and aligned to the user's satisfaction, the three-dimensional model is built. First, the backbone atom positions of the template are averaged and atoms that significantly deviate from the target are eliminated. Insertions and deletions are managed by assembly of a collection of fragments that could be compatible to fit the neighboring stems, and one is selected using a scoring scheme based on force field energy, steric hindrance, and favorable interactions (ie. hydrogen bonds). After the model backbone has been established, side chain modeling is carried out based on weighted positions of corresponding residues in the template. Side chain positions are assigned by a scoring function based on favorable interactions (ie. hydrogen and disulfide bonds) and unfavorably close contacts [6]. Finally the model is assessed for energy minimization via a GROMOS96 force field in order to detect aspects of the structure with conformational errors [6, 9]. When interpreting homology models it is important to take these methods into account. Clearly backbone and side chain fidelity are more constrained in regions that have secondary structure while loops are far more variable and difficult to predict and thus should be considered less reliable.

The final step in generating a homology model is to evaluate the quality of the model. Initially, the target sequence can be analyzed via InterPro and PsiPred [10, 11], which predict conserved domains and DISOPRED [12], which predicts disordered

regions in the protein. Assuming the model is consistent with these predictions, the stereochemical likelihood of the model can be appraised by programs such as PROCHECK that assess amino acid conformations that may deviate from expected values (ie. bond angle calculations and Ramachandran plots) [13]. QMEAN is used to generate a composite scoring function for total model quality [14]. More specifically, DFIRE calculates a protein conformation free energy score [15] and ANOLEA (Atomic Non-Local Environment Assessment) assesses the mean force potential to determine the energy environment for each amino acid [7]. All of these analyses are used to judge the fit of the model to the target; more specifically they can be used to identify regions of the model that have higher fidelity than others and can therefore be assessed with more faith.

ACF1.

In order to examine the possible secondary and tertiary structures of ACF, we began by homology modeling ACF1. Amino acids C56 through D134 were submitted to SWISS-MODEL to identify potential templates through multiple BLAST search mechanisms. A large number of templates were identified, but few were more than 30% identical to the sequence of ACF1. The five templates with highest identity were selected and aligned with ACF1 both by SWISS-MODEL and ClustalW2 [8] and the alignment with the highest homology used to generate a model of ACF1's structure (Figure 4.2). It should be noted that models were generated for each template based on both the SWISS-MODEL and ClustalW2 alignments; there were very few differences in the models based on the two different alignments, which resulted in negligible changes in the final three-dimensional structures. All the proteins identified as satisfactory templates are RRM-

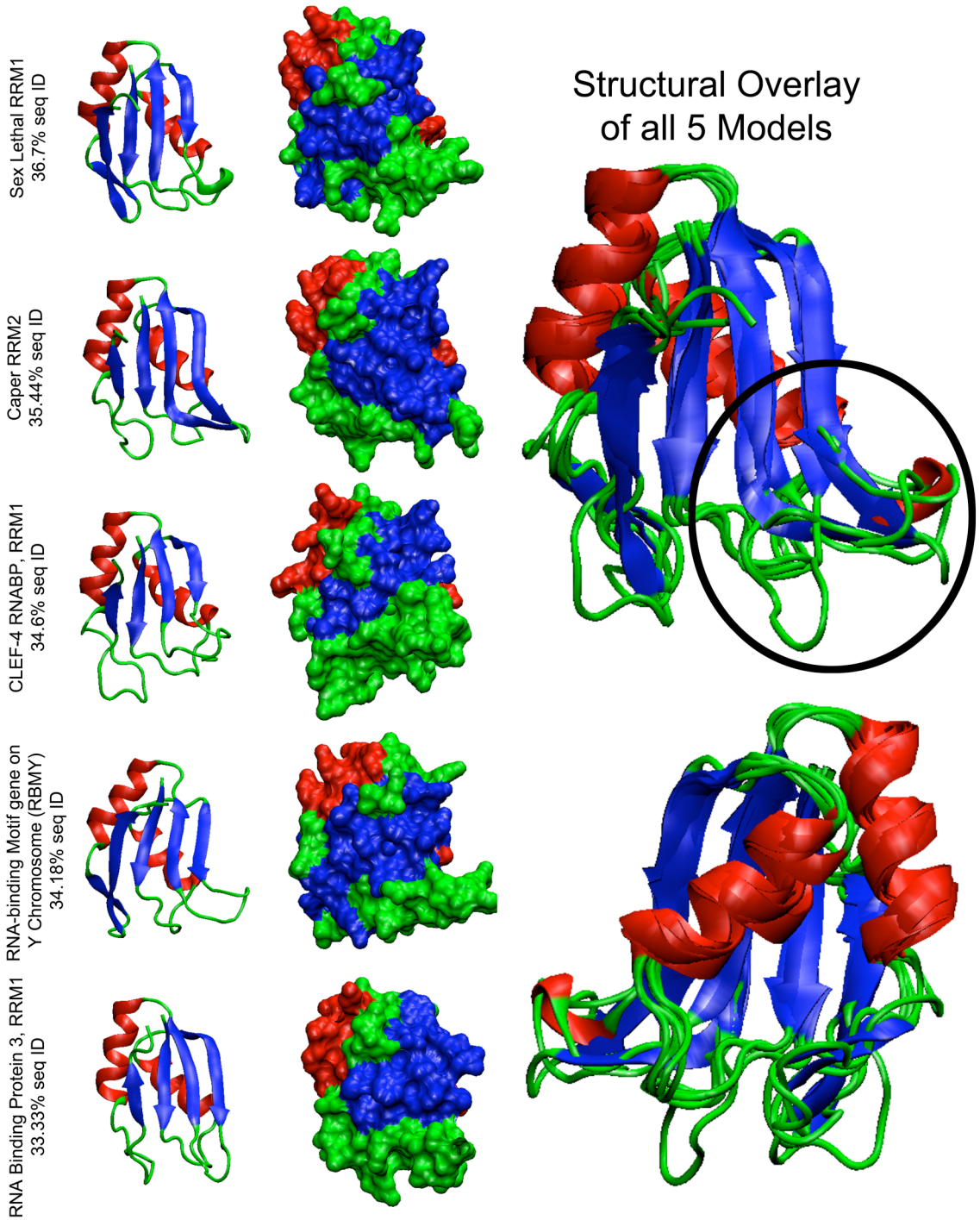


Figure 4.2 - Homology models of ACF1. Left) Five different homology models in ribbon and space-filling representations. Models generated from five different template structures: Sex Lethal (pdb: 1b7f), Caper (pdb: 2jrs), CLEF-4 (pdb: 2dgp), RBMY (pdb: 2fy1), RBP3 (pdb: 2dnq). Right) The five models are overlaid using MultiSeq structural alignment. Circle indicates location of loop 3.

containing RNA binding proteins. Specifically, they are: Sex Lethal, Caper, CLEF-4, RNA Binding Protein 30, and RBMY. All are implicated in alternative splicing; RBMY additionally regulates spermatogenesis.

All the potential models of ACF1 exhibit the classic RRM global structure of a four-stand anti-parallel β-sheet supported from behind by two alpha-helices (Figure 4.2, ribbon models). As is evident in the space filling representations, in all cases the beta-sheet is accessible on the surface of the protein. The β-sheet is the most common RNA binding site for RRMs. To more directly compare these possible models, the five structures were overlaid using MultiSeq [16]. It is clear that while each of the models is unique, they exhibit striking similarity both in global organization as well as definition of the secondary structure.

Though these models display the predicted global arrangement of an RRM, it is important to assess their quality. Therefore, we further examined the quality of the Sex Lethal-based model (that with the highest sequence identity to ACF1) using the parameters described above. The ANOLEA calculations determined that only five amino acids within the domain had unfavorable energy environments; all of these amino acids are in loop 3 (between β-strand 2 and β-strand 3). This is not surprising, as loops can be the most difficult regions to model due to the lack of structural constraints. This area is of particular interest because in multiple RRM proteins, loop 3 makes direct contact with the RNA and has been associated with sequence specificity (indicated by a circle on the multi-model overlay) [17, 18]. Therefore, though this region may be the least reliable according to the models, it will be one of the most important to define in experimental determination of ACF's structure. Further analysis revealed this model has a DFire free

energy score of -92.37 suggesting it is thermodynamically favorable [15]. The QMEAN analysis of the global model gave a score of 0.802 (on a scale of 0-1 with 1 being the most ideal structure) and a z-score of 0.485 (a weighted score in which the more negative the value, the less reliable the model) [14]. Finally PROCHECK analysis examined the conformational constraints of each amino acid [13]. Main-chain bond lengths and bond angles were over 96% within limits and only one residue fell in disallowed regions of the Ramachandran plot (specifically: R125, in loop 5), which once again is not surprising due to the difficulty in accurately modeling unstructured regions of the protein. While this structure is only a model and therefore all interpretations are heavily subject to experimental verification, it is enticing to estimate this structure to be reflective of the native fold of ACF1 *in vivo*.

ACF2.

While ACF RRMs 1 and 3 were expressed as recombinant proteins, no construct containing RRM2 was successfully expressed in *E. coli* (including RRM2, RRMs 1+2, and RRMs 2+3). As such, we questioned why expression of RRM2 appeared to confer some degree of protein instability or insolubility in bacterial cells that prevented recombinant expression. Therefore, we examined the structures of ACF380, ACF1, and ACF3 by circular dichroism to assess the molar ellipticity of each of the recombinant protein constructs we generated (Figure 4.3A). Molar ellipticity is a measurement of the amount of secondary structure per protein molecule. Therefore, the molar ellipticity of the whole (i.e. ACF380) should be the sum of its parts (i.e. ACF1, ACF2, ACF3). A large disordered region, such as failure of ACF2 to fold, may explain the challenge in bacterial

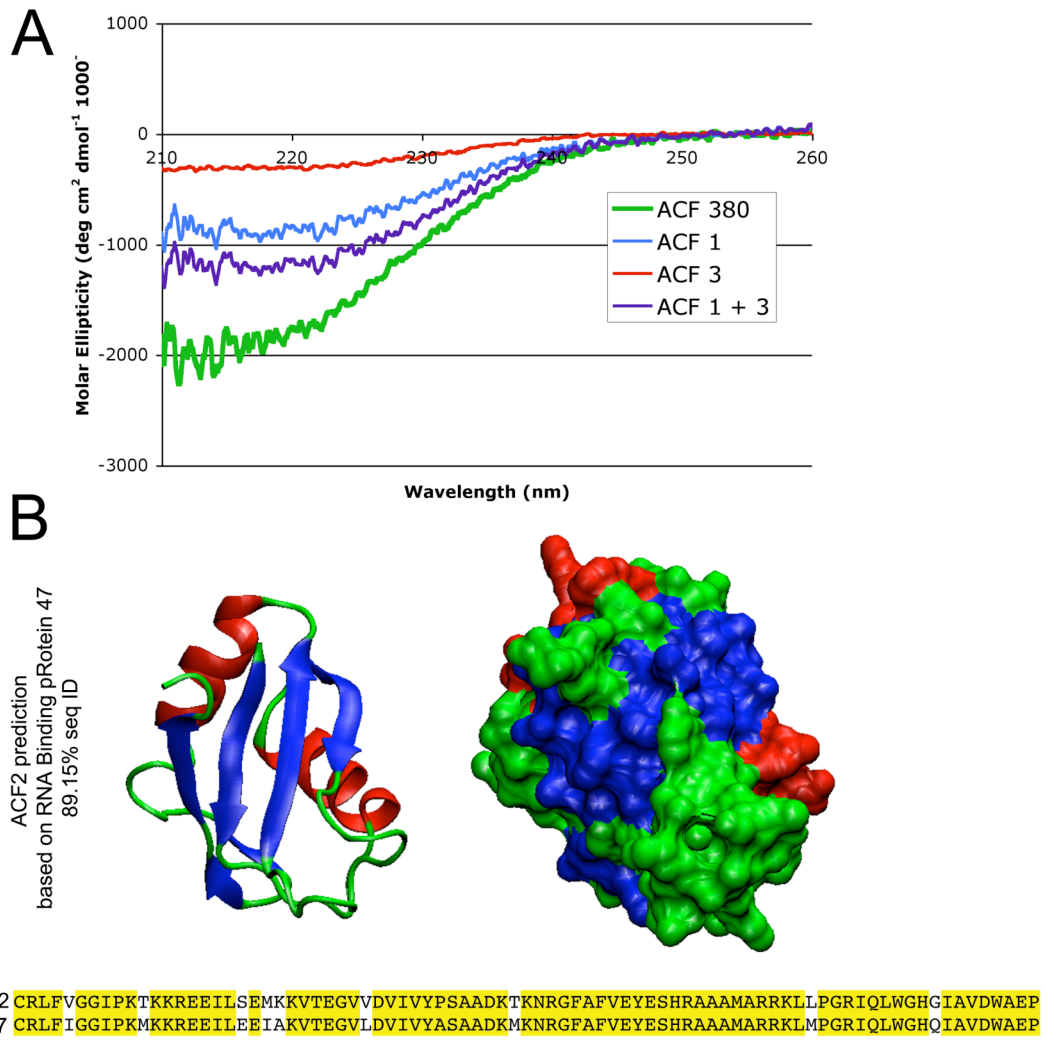


Figure 4.3 - ACF2. A) UV CD spectra of ACF380, ACF1, and ACF3 plotted as molar ellipticity as a function of wavelength (nm). ACF1+3 represents sum of ACF1 and ACF3 data. B) Homology model in ribbon (left) and space-filling (right) representations of ACF2 generated from RNA Binding Protein 47 template (pdb: 2dis). Alignment of ACF2 and RBP47 with identical residues highlighted in yellow.

expression of these recombinant constructs. If ACF2 fails to fold into the predicted RRM in the context of ACF380, the molar ellipticity of ACF380 should be equivalent to the sum of ACF1+3. However, we see that ACF1+3 has significantly less signal than ACF380. The difference is similar to the amount of structure observed in ACF1. This suggests that ACF2 does adopt a secondary structure with signal approximately equal to that of an RRM, as predicted by sequence homology.

Homology modeling of ACF RRM2 was different than that of RRM1. Upon searching for possible templates, one sequence dramatically stuck out. RNA Binding Protein 47 (RBP47) exhibits 89.15% sequence identity with ACF2 (Figure 4.2B), while no other sequences examined by any of the search methods used by SWISS-MODEL had sequence identity greater than 25%. Alignment of ACF2 and RBP47 was optimized by ClustalW2 [8] and modeled (Figure 4.3B). Once again this model exhibits the predicted global arrangement of an RRM and the space-filling representation reveals the exposed β-sheet for potential protein:RNA interaction.

Analysis of this model revealed very few residues with positive ANOLEA scores (three in loop 3 and four in α4), but all with very low absolute values. The DFire free energy score of -92.85 suggests the global conformation is favorable, and the QMEAN score of 0.766 and z-score of 0.212 suggests high model reliability. According to PROCHECK analysis, all residues are within allowable regions on the Ramachandran plot and main-chain bond lengths and bond angles are 100% and 98.9% within limits, respectively. These data, together with the significant sequence similarity between ACF2 and RBP47 suggest this is the optimal model for prediction of ACF2's conformation; however, this structure still must be experimentally confirmed.

ACF12.

In vitro binding assays described in previous chapters indicate that both RRMs 1 and 2 contact RNA and likely work synergistically to result in ACF:RNA interaction. This dual binding of the two RRMs in tandem makes it even more important to understand the structural relationship between these two independent domains. Therefore, we undertook SWISS-MODEL homology modeling of amino acids 56 to 218 (encompassing RRMs 1 and 2 and the linker between them) in order to predict the possible interface between the two domains.

After searching for potential templates, it became clear that this model would need to be built from a template with significantly less sequence identity than those found for ACF1 or ACF2. Three templates were identified: Sex Lethal, HRP-1, and HuC (a homolog of HuR and HuD), all with 25-28% sequence identity (Figure 4.4). All three template proteins contain at least 2 RRMs in tandem that work synergistically to bind U-rich RNA sequences, much in the same way we hypothesize ACF functions. All three models exhibit overall similar topology, though the HRP-1-based model forms a helix in the interdomain linker (see arrow in HRP-1 ribbon model, Figure 4.4). When overlaid, the models exhibit similar geometry and domain orientation (Figure 4.5), indicating that alignment with any of these three molecules results in a similar global structure.

Quality assessment of the models lead us to believe these structures can be useful but not as reliable as those obtained for the individual domains. The primary metrics used to assess energy and model reliability trended positively, but were less convincing than those of the ACF1 and ACF2 models (Table 4.1). In all cases, the DFire free energy score

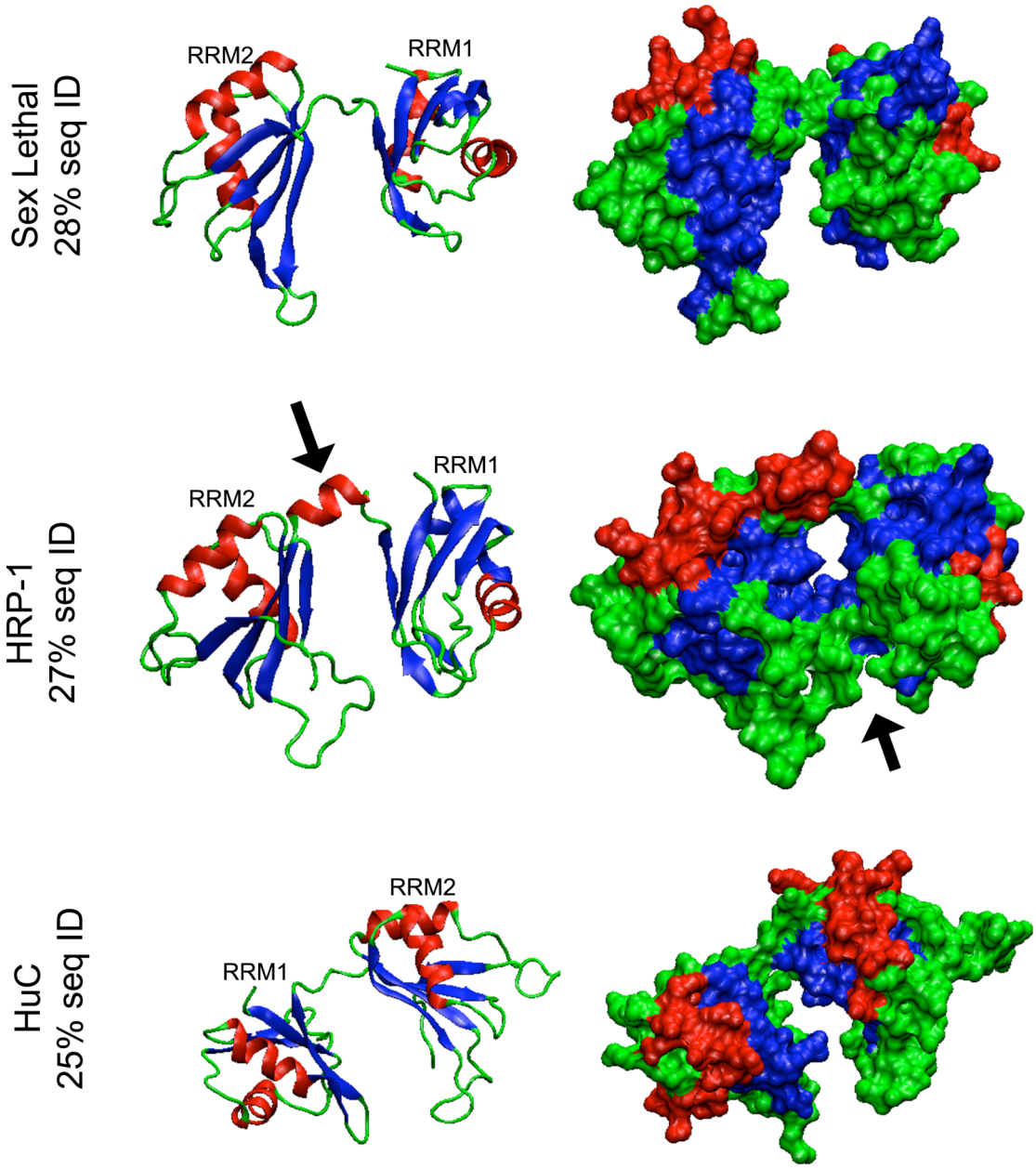


Figure 4.4 - Homology models of ACF12. Models were generated from three templates: Sex Lethal (pdb: 1b7f), HRP-1 (pdb: 2cjk), and HuC (pdb: 1fnx). Ribbon (left) and space-filling (right) representations of possible ACF12 structures. Left arrow indicates predicted helix in interdomain linker. Right arrow indicates predicted interdomain contact.

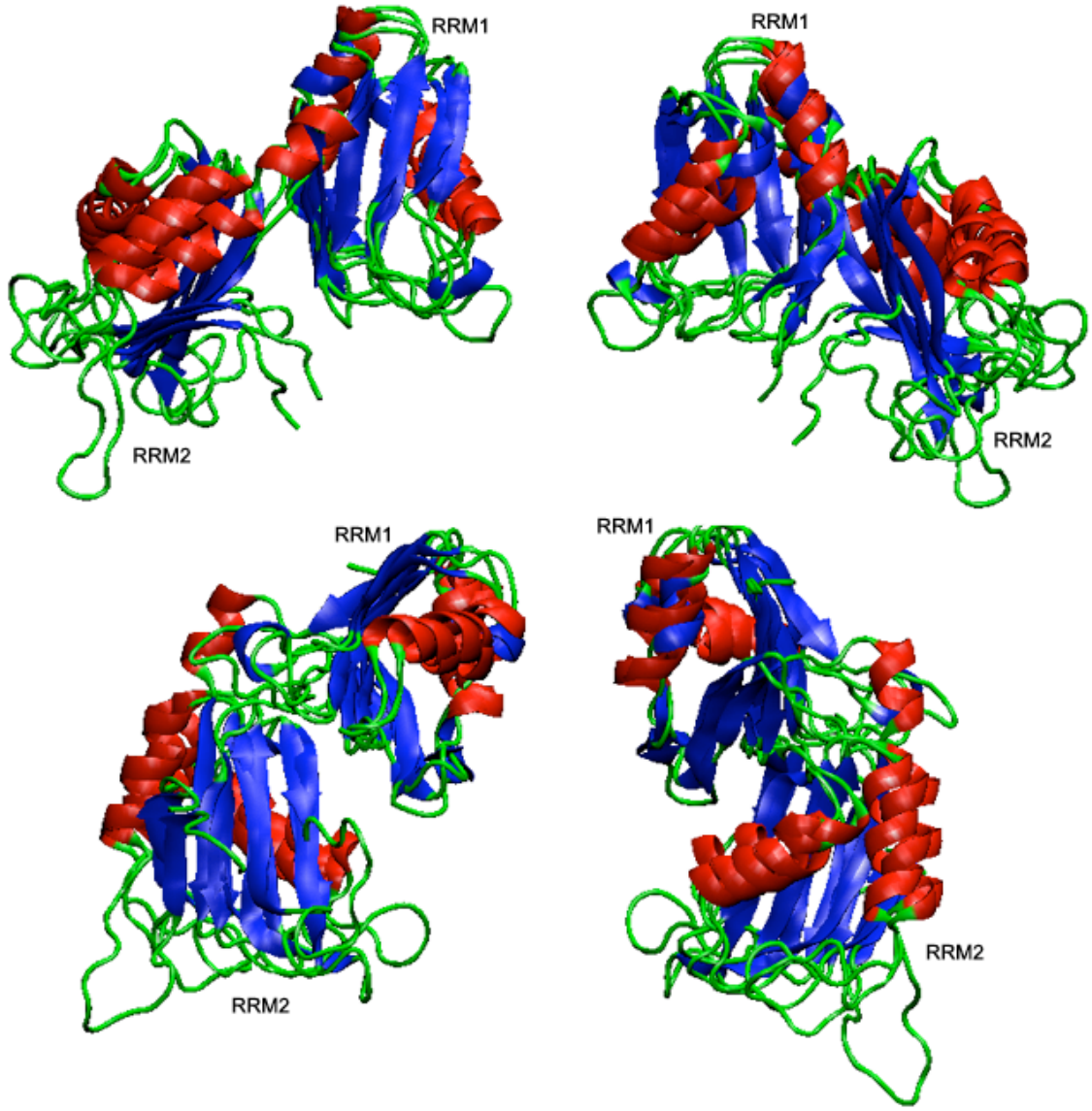


Figure 4.5 - Overlay of ACF12 models. The three ACF12 models described in Figure 4.4 have been aligned by MultiSeq. Four different vantage points are shown to accentuate the front and back views of each RRM.

Table 3.1 - ACF12 Model Quality. The model quality metrics for each of the three possible ACF12 models.

Metric	Sex Lethal Model	HRP1 Model	HuC Model
DFire	-176.42	-158.77	-158.77
QMEAN	0.73	0.576	0.586
z-score	-0.138	-1.698	-1.598
Ramachandran Disallowed Residues	1	1	2
Main-Chain bond angle	94.5%	95.3%	93.6%
Main-Chain bond length	98.1%	96.8%	98.8%

was low enough to consider the conformation favorable. The QMEAN score of the Sex Lethal-based model indicated reliability while the other two models were more questionable with lower QMEAN scores and z-scores that were more negative. All three models were only satisfactory in terms of the Ramachandran plot and main-chain bond lengths and angles. These data suggest to us that details of these models are not reliable. Thus, it is even more important that the structure of ACF be experimentally determined in order to understand the mechanism by which it binds RNA and to observe how various RNA sequences may impact its mode of interaction.

Though these models must not be interpreted strictly, there are still observations that can be made regarding the interaction of RRMs 1 and 2. In each of the models, the two RRMs are oriented in such a way that the β-sheets essentially face each other. As indicated in the space-filling representations, the two domains are in very close proximity, close enough for potential physical contact (see arrow in HRP-1 space-filling representation, Figure 4.4) or for hydrogen bonds, as observed in the experimentally determined structures of Sex Lethal and HRP-1 [19]. However, it is particularly important in this case to remember that the linker is flexible. The separation of the two domains is likely variable in solution and certainly needs to be experimentally determined in order to interpret these observations accurately.

The linker is an interesting aspect of these tandem RRM domains (Figure 4.6A). In all three template proteins (Sex Lethal, HRP-1, and HuC), their RNA substrates have been observed to make contact with the linker in a specific manner. In solution NMR studies of Sex Lethal in the presence or absence of U-Rich RNA, the most significant chemical shifts were found in the linker, the RNP motifs, and one residue in loop 3 [20].

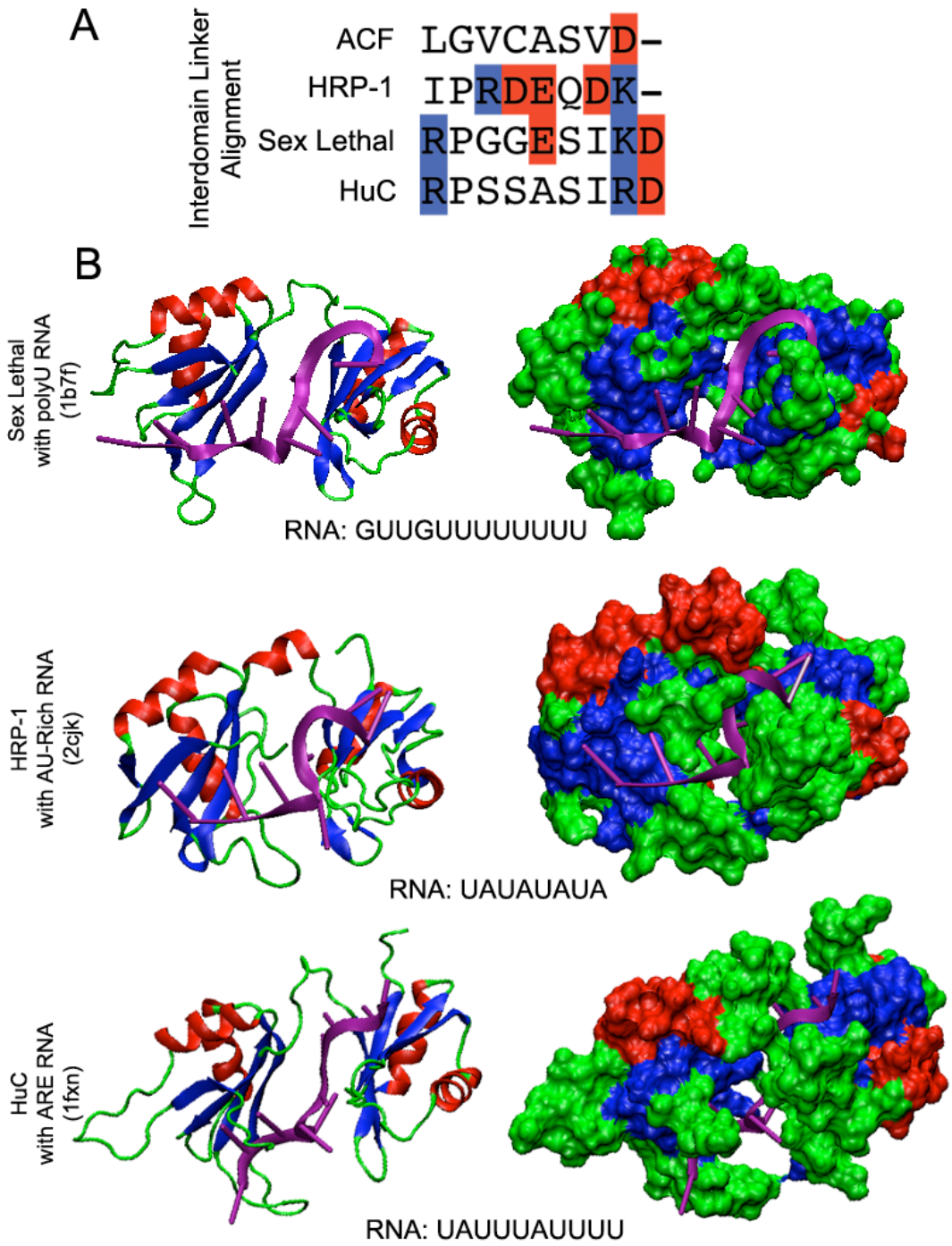


Figure 4.6 - ACF Homologs. A) Alignment of the linker region of ACF12 (as defined by the HRP-1-based model) with the linkers of the template molecules. Positive residues are highlighted in blue and negative residues are highlighted in red. B) Experimentally determined structures of the three homologous proteins used as templates for ACF12 modeling. Target RNA is represented in purple in both the ribbon (left) and space-filling (right) representations with primary RNA sequence listed below structures.

This is supported by biochemical data suggesting that elimination of the linker drastically reduces Sex Lethal's affinity for U-rich RNA [21]. Crystal structure analysis of Sex Lethal indicated that the structure of the linker allowed no interdomain contact between the RRMs in the crystallized conformation, but suggests that upon RNA binding there is re-organization of both the RNA and the tandem binding surfaces to cause interdomain interactions that contribute to increased affinity and specificity [22].

Interestingly, the linker of HRP-1 is also involved in RNA contact, as solution NMR data reveal that upon addition of RNA the loop adopts a helical structure that is stabilized by salt bridge interactions between charged residues in the linker [19]. While the RNA nucleotides appear to be bound by RRMs 1 and 2 of HRP-1, the linker residues actively participate in the recognition of the target. The HRP-1-based model of ACF12 proposes a similar helix in the interdomain linker, however the linker region of ACF has far fewer charged residues (only one Asp) in the linker, suggesting that the salt-bridge stabilization would not be likely (Figure 4.6A).

All three template structures were solved in the presence of RNA. Therefore, while the ACF12 structures are only models, it is informative to look at the interaction of three highly homologous molecules that all interact with similar AU-Rich RNA targets (Figure 4.6B). In Figure 4.6, the experimentally determined structures of Sex Lethal, HRP-1, and HuC are displayed bound to RNA. It is interesting that all three targets bind ssRNA, as also observed with ACF (Chapter 3). In all cases, the RNA seems to thread through the tandem β-sheets and interact with a number of protein sites including the interdomain linker and loop 3 of one or both RRMs.

To chart a hypothetical path of the RNA on ACF, we have examined the electrostatic surface potential of the models of ACF1, ACF2, ACF3, and ACF12 (in the case of ACF1 and ACF12, the Sex Lethal-based models were used as these were the models with the highest sequence identity and quality assessment). Using Adaptive Poisson-Boltzman Software (APBS), the structures of each model were evaluated to estimate the charge potential on the surface in 100 mM salt (which is equivalent to that used in *in vitro* binding assays) [23]. On each representation, blue indicates positive charge, which would be expected to electrostatically interact with negatively charged RNA, and red indicates negative charge (Figure 4.7). As expected, the β-sheet of all three RRMs has at least a weak positive charge. The ACF12 model is particularly interesting, as it suggests loop 3 of RRM2 has high positive charge density, suggesting RNA may interact with this loop. While these potentials are merely predictions since they are based on model structures, it is clear that the likely RNA binding sites on all these domains exhibit charge potentials that are conducive to RNA binding.

Conclusion.

It is useful to examine these ACF12 models in light of the ACF binding activity observed in Chapter 3. We believe ACF's preference for ssRNA is consistent with the predicted relative orientation of RRMs 1 and 2, as all three template molecules bind ssRNA. We examined the interaction of RRM2 with Cox-2 RNA through tryptophan fluorescence. In Figure 4.8 we have highlighted the positions of the tryptophans found in RRM2 on the modeled structures of both ACF2 and ACF12. In both cases, Trp 215 is predicted to form the edge of β-strand 4, but is likely oriented to the interior of the

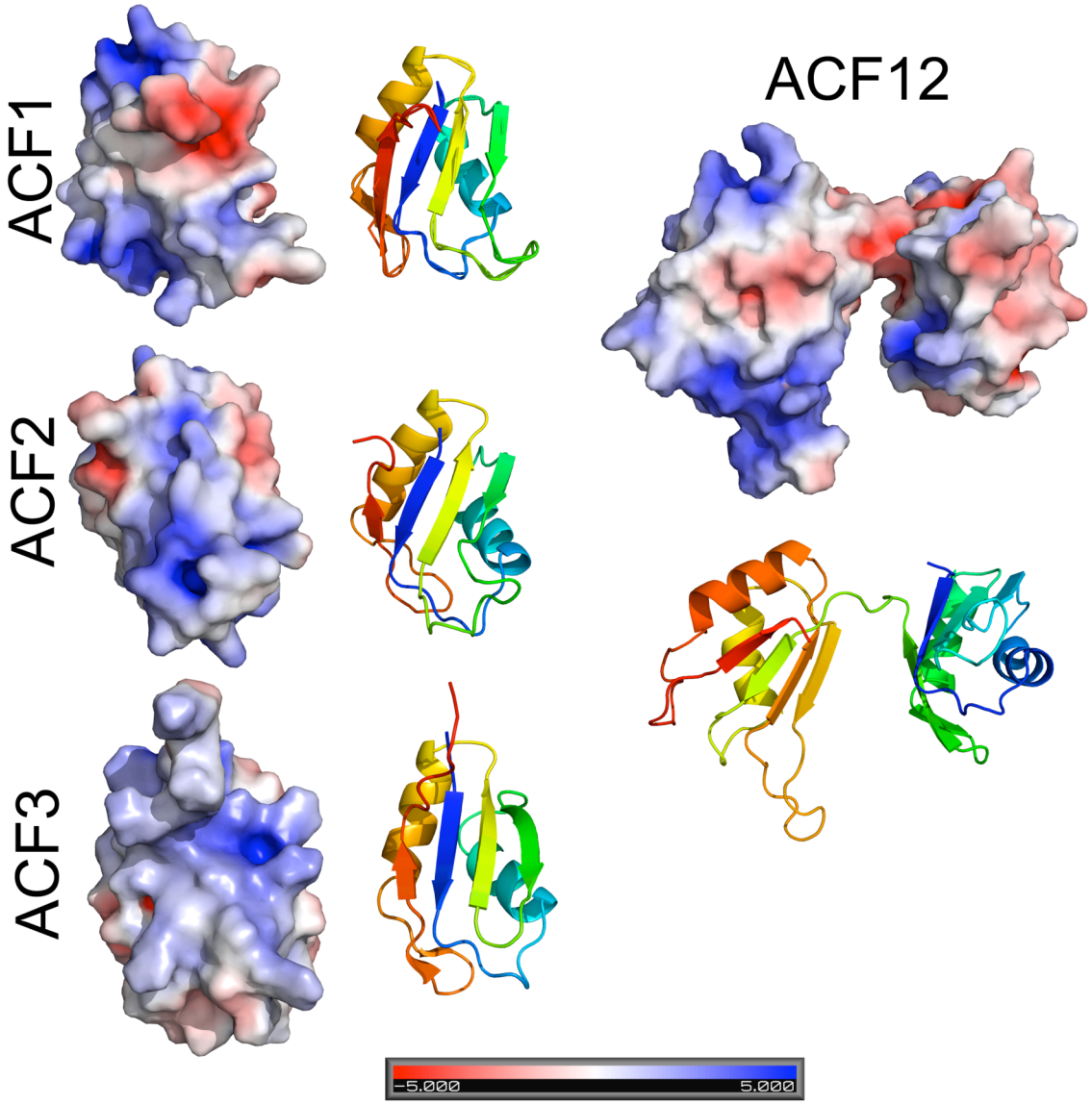


Figure 4.7 - Electrostatic surface potentials of ACF models. Models of ACF1, ACF2, and ACF12 along with the structure of ACF3 shown in space-filling representations with *predicted* electrostatic surface potentials indicated by red to blue spectrum. Adjacent ribbon representations indicate what aspect of the molecule is shown.

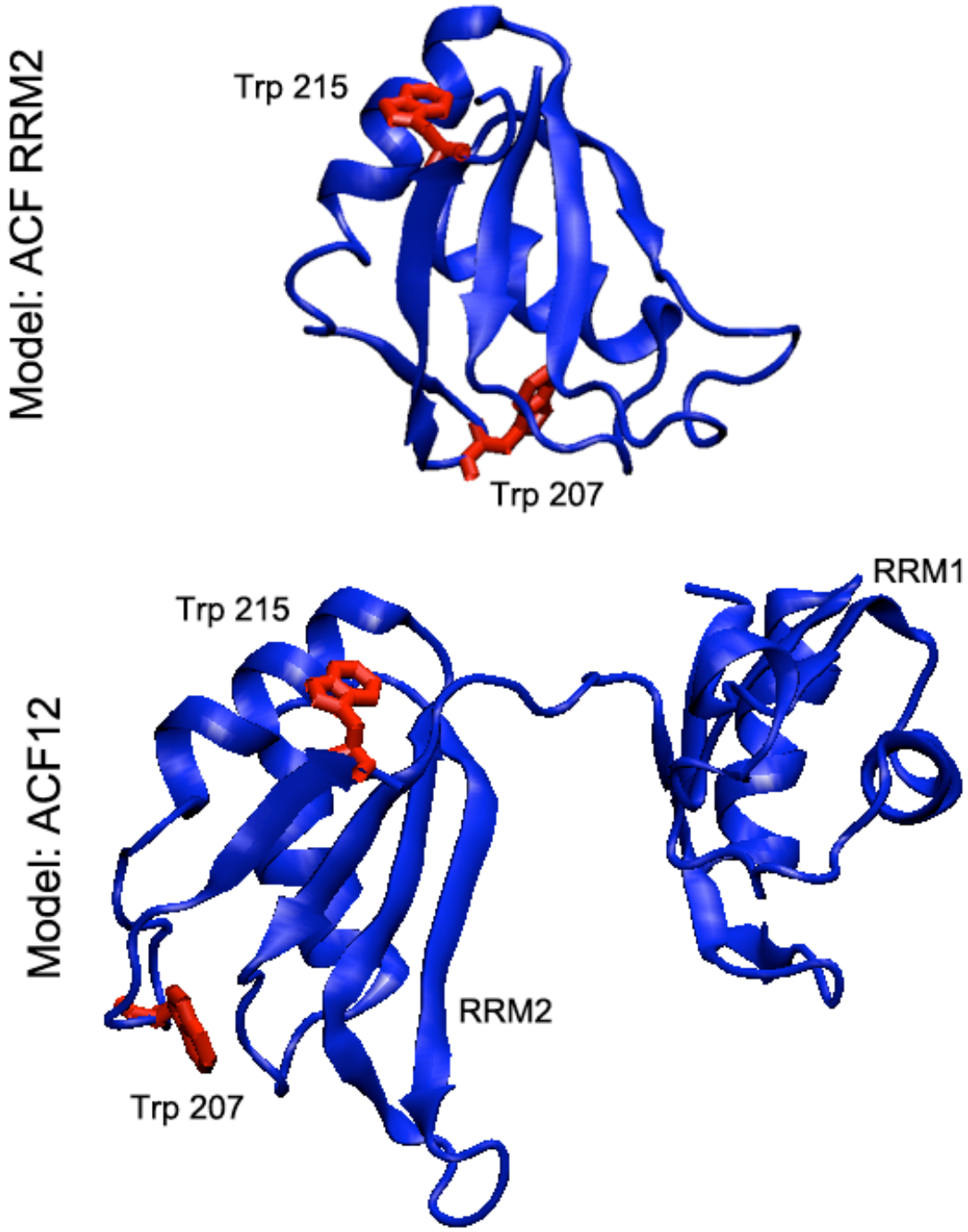


Figure 4.8 - Tryptophans in ACF. The homology models of ACF2 and ACF12 (Sex Lethal-based model) are shown in blue with Trp 207 and Trp 215 bases in red.

protein and not RNA-accessible, while Trp 207 is found in loop 5. Of course it is important to remember these are only models and the tryptophans may be oriented in a different fashion; however we believe this orientation of Trp 207 is consistent with the tryptophan quenching we observed upon Cox-2 binding. It is possible that as the ssRNA binds the β-sheet, it interacts with the tryptophan in loop 5. In HRP-1, an aromatic residue in the loop has been shown to stack with an RNA base to contribute to both affinity and specificity [19]. This would account for the quenching observed in ACF380 upon addition of Cox-2 RNA. Alternatively, the quenching may be due to a domino effect of protein rearrangement. As mentioned before, it is very common for loop 3 to participate in RNA interaction. It has also been observed that loops 3, 1, and 5 pack against each other. We hypothesize that Cox-2 RNA interaction with RRM2 loop 3 may cause subsequent shifts in loops 1 and 5 that could result in a net change of the local environment surrounding Trp 207 and thus the observed quenching.

While homology modeling has significant limitations, we feel that the theoretical data presented here allows intriguing interpretation of the properties of ACF as an RNA binding protein. It is reasonable to hypothesize that ACF RRMs 1 and 2 adopt a tandem orientation in which target RNAs must thread between the two β-sheets that face each other. This may account for ACF's apparent preference for ssRNA in some targets. However, as discussed in previous chapters, at least one of ACF's targets (ApoB) is believed to adopt a stem-loop conformation. A stem-loop would clearly not be an optimal target for ACF binding given the *predicted* orientation of RRMs 1 and 2. This may be the reason ACF380 has a higher *in vitro* affinity for IL-6 and Cox-2 than ApoB RNA (Chapter 3). This is also consistent with ACF1's lack of participation in ApoB mRNA

binding (Chapter 3, published data, [24]); it is possible that ACF cannot bind the ApoB stem-loop RNA in the same manner that it binds ssRNA targets and therefore protein re-arrangements (induced by RNA recognition or protein-protein interactions) may result in ApoB binding by only RRM2.

ACF presents a particularly curious behavior as an RNA binding protein. It has been extensively characterized as part of the ApoB mRNA editing complex and we have shown that it regulates mRNA stability of multiple targets in a differential manner. Our *in vitro* binding data indicate that ACF interacts with each of its targets by a different mechanism. This is also consistent with our homology modeling data, which indicate ACF could not accommodate ssRNA (such as Cox-2) and stem-loop RNA (such as ApoB) in the same manner. Therefore, it is clear that the next step in understanding the role of ACF in regulation of multiple RNA targets both by RNA editing and differential control of RNA stability is to experimentally characterize the structure of ACF both alone and in combination with each of its RNA targets.

REFERENCES

1 K. Arnold, L. Bordoli, J. Kopp and T. Schwede. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling, Bioinformatics. **22**, 195-201

2 L. Bordoli, F. Kiefer, K. Arnold, P. Benkert, J. Battey and T. Schwede. (2009) Protein structure homology modeling using SWISS-MODEL workspace, Nat Protoc. **4**, 1-13

3 T. Nagata, Y. Muto, M. Inoue, T. Kigawa, T. Terada, M. Shirouzu and S. Yokoyama. (2005) Solution structure of the RNA recognition motif of human APOBEC-1 complementation factor, ACF, Protein Database, pdb.org.

4 F. Kiefer, K. Arnold, M. Kunzli, L. Bordoli and T. Schwede. (2009) The SWISS-MODEL Repository and associated resources, Nucleic Acids Res. **37**, D387-392

5 J. Kopp and T. Schwede. (2004) The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models, Nucleic Acids Res. **32**, D230-234

6 T. Schwede, J. Kopp, N. Guex and M. C. Peitsch. (2003) SWISS-MODEL: An automated protein homology-modeling server, Nucleic Acids Res. **31**, 3381-3385

7 F. Melo and E. Feytmans. (1998) Assessing protein structures with a non-local atomic interaction energy, J Mol Biol. **277**, 1141-1152

8 R. Chenna, H. Sugawara, T. Koike, R. Lopez, T. J. Gibson, D. G. Higgins and J. D. Thompson. (2003) Multiple sequence alignment with the Clustal series of programs, Nucleic Acids Res. **31**, 3497-3500

9 W. van Gunsteren. (1996) Biomolecular Simulations: The GROMOS96 Manual and User Guide., VdF Hochschulverlag ETHZ.

10 E. M. Zdobnov and R. Apweiler. (2001) InterProScan--an integration platform for the signature-recognition methods in InterPro, Bioinformatics. **17**, 847-848

11 D. T. Jones. (1999) Protein secondary structure prediction based on position-specific scoring matrices, J Mol Biol. **292**, 195-202

12 J. J. Ward, J. S. Sodhi, L. J. McGuffin, B. F. Buxton and D. T. Jones. (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life, J Mol Biol. **337**, 635-645

13 R. A. Laskowski, D. S. Moss and J. M. Thornton. (1993) Main-chain bond lengths and bond angles in protein structures, J Mol Biol. **231**, 1049-1067

14 P. Benkert, T. Schwede and S. C. Tosatto. (2009) QMEANclust: estimation of protein model quality by combining a composite scoring function with structural density information, BMC Struct Biol. **9**, 35

15 H. Zhou and Y. Zhou. (2002) Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction, Protein Sci. **11**, 2714-2726

16 E. Roberts, J. Eargle, D. Wright and Z. Luthey-Schulten. (2006) MultiSeq: unifying sequence and structure data for evolutionary analysis, BMC Bioinformatics. **7**, 382

17 D. Scherly, W. Boelens, N. A. Dathan, W. J. van Venrooij and I. W. Mattaj. (1990) Major determinants of the specificity of interaction between small nuclear ribonucleoproteins U1A and U2B" and their cognate RNAs, Nature. **345**, 502-506

18 S. R. Price, P. R. Evans and K. Nagai. (1998) Crystal structure of the spliceosomal U2B"-U2A' protein complex bound to a fragment of U2 small nuclear RNA, Nature. **394**, 645-650
19 J. M. Perez-Canadillas. (2006) Grabbing the message: structural basis of mRNA 3'UTR recognition by Hrp1, EMBO J. **25**, 3167-3178
20 A. L. Lee, B. F. Volkman, S. A. Robertson, D. Z. Rudner, D. A. Barbash, T. W. Cline, R. Kanaar, D. C. Rio and D. E. Wemmer. (1997) Chemical shift mapping of the RNA-binding interface of the multiple-RBD protein sex-lethal, Biochemistry. **36**, 14306-14317
21 M. Samuels, G. Deshpande and P. Schedl. (1998) Activities of the Sex-lethal protein in RNA binding and protein:protein interactions, Nucleic Acids Res. **26**, 2625-2637
22 S. M. Crowder, R. Kanaar, D. C. Rio and T. Alber. (1999) Absence of interdomain contacts in the crystal structure of the RNA recognition motifs of Sex-lethal, Proc Natl Acad Sci U S A. **96**, 4892-4897
23 N. A. Baker, D. Sept, S. Joseph, M. J. Holst and J. A. McCammon. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome, Proc Natl Acad Sci U S A. **98**, 10037-10041
24 V. Blanc, J. O. Henderson, S. Kennedy and N. O. Davidson. (2001) Mutagenesis of apobec-1 complementation factor reveals distinct domains that modulate RNA binding, protein-protein interaction with apobec-1, and complementation of C to U RNA-editing activity, J Biol Chem. **276**, 46386-46393