

RNA-Binding Proteins which Affect mRNA Translation or Stability

Jackie Wilce¹, Peter Leedman² and Matthew Wilce³

¹Department of Biochemistry and Molecular Biology and the Department of Chemistry, University of Western Australia, Perth, WA 6009

²Laboratory for Cancer Medicine, Western Australian Institute for Medical Research, University Department of Medicine, University of Western Australia, Perth, WA 6009

³Crystallography Centre and Department of Pharmacology, University of Western Australia, Perth, WA 6009

The Wilce and Leedman Laboratories have joined forces at the University of Western Australia to study the biophysical features of protein/RNA interactions involved in the regulation of gene expression. In particular, those factors that determine the translational efficiency and stability of specific mRNAs are of interest for their potential as targets in modulation of protein production.

mRNA stability

The regulation of gene expression can take place at many stages subsequent to transcriptional activation. In eukaryotic systems the messenger RNA (mRNA) is subject to processing, nucleocytoplasmic transport, translation and degradation. Each of these points serves as a checkpoint at which the ultimate production of encoded protein can be regulated. Studies in the past two decades have shown that the modulation of mRNA stability plays an important role in regulating gene expression (1,2). In particular, factors that affect mRNA stability dictate whether the message half life is minutes or hours. This can underlie 1000-fold differences in the mRNA available for translation, and thus the amount of protein produced (3).

Early alterations of the mRNA transcript are vital to its translational efficiency and stability. These include capping of the 5' terminal nucleotide and polyadenylation of the 3' end. The cap, which is an N7-methylated guanosine triphosphate, helps initiate pre-mRNA splicing, nucleocytoplasmic transport, mRNA end formation and translation, and protects the transcript from 5' to 3' exonucleases. Only when the 5' cap is removed can degradation occur in a 5' to 3' direction. The poly-A tail is bound by the poly-A tail-binding protein (PABP) that assists with the initiation of translation. The poly-A tail also temporarily protects the coding region of the transcript from 3' to 5' exonucleases.

mRNA stability is also dictated by the presence of specific *cis*-acting elements (e.g. AU-rich elements (AREs)), or other instability elements, most often located in their 3' untranslated regions (UTRs) (4). AREs are generally 50 to 150 bases long and contain either one to three scattered copies of the AUUUA pentamer (class I), clustered and overlapped AUUUA repeats (class II), or contain long continuous U-rich regions (Class III). AREs have been found to direct mRNA degradation and thus play an important role in the regulation of gene expression. A number of proteins have been identified which specifically bind AREs and directly affect the stability of the mRNA. Likewise, the binding of proteins to poly-C motifs in the 3'

UTR affects mRNA stability and translation rate (5). Such post-transcriptional events are proving to be of enormous importance in the rapid flux of gene expression in response to signalling events, and have evoked a great interest in studies of RNA-protein interactions (6).

Proteins that bind RNA

The number and variety of RNA-binding proteins reported in the literature are rapidly expanding. Some show low binding specificity and are likely to have general packing interactions, whilst others have high RNA-binding specificity, suggesting defined roles in gene regulation. RNA-binding proteins, like their DNA-binding protein counterparts, tend to be modular in structure with conserved RNA-binding motifs alongside other functional domains. Four common motifs that have been identified and structurally characterised are the RNA recognition motif (RMM; also referred to as RNP), the RGG box, zinc fingers and the KH (hnRNP K Homology) domain (7). These motifs often occur as tandem repeats within a single protein, however information as to their structural juxtaposition is not available in most cases.

HuR and CPs both bind androgen receptor mRNA

We are currently investigating factors that dictate the stability of androgen receptor (AR) mRNA. The AR, a member of the nuclear receptor superfamily of transcription factors, plays a central role in the proliferation of prostate cancer cells and is a major target for novel therapeutics. Defining the mechanisms that modulate AR gene expression would be of immense value for the development of novel prostate cancer therapies.

We have previously shown that AR mRNA stability is an important determinant of AR gene expression in human cancer cells (breast and prostate) (6). In recent studies to elucidate the mechanism underlying the regulation of AR mRNA decay, we identified a highly conserved UC-rich region in the 3' UTR of AR mRNA. This region contains several C-(U)_n-C motifs and a CCCUCCC poly-C binding protein (CP) motif and thus constitutes a possible regulatory element on the mRNA. We showed that this sequence, in fact, modifies expression of a luciferase reporter gene.

Subsequent work established that several cytoplasmic and nuclear RNA-binding proteins from human prostate cancer cells target this sequence. These include HuR, a

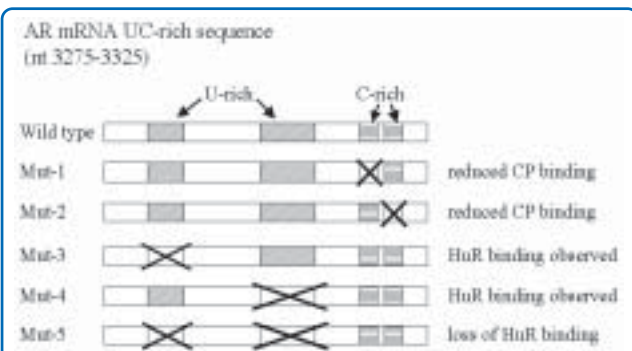


Fig. 1. Depiction of the AR mRNA UC-rich sequence and the effect of mutated regions on HuR and CP binding.

The major U-rich and C-rich regions within AR mRNA (nt 3275-3325 GenBank Accession No: M20132) are highlighted by hashed boxes. The regions that were targeted for mutation are indicated by a cross, and the effect of the mutation on either HuR or CP binding is listed. Both U-rich regions are involved in HuR binding and both C-rich regions are required for CP binding.

member of the *Elav/Hu* family of RNA-binding proteins involved in the stabilisation of several mRNAs (9), and the poly-C binding proteins, CP1 and CP2 (also referred to as hnRNPs E1 and E2 or PCBPs 1 and 2), which are also implicated in the control of mRNA turnover and translation (10).

Mutational analysis of the AR mRNA was used to determine the binding region for each of the proteins. HuR binding was found to be dependent on two U-rich regions within a defined 51-base pair region of the AR mRNA 3' UTR. Both CP1 and CP2 binding was found to be dependent on the CCCUCCC motif, with mutation of either CCC motif causing reduced binding (Fig. 1).

Protein binding studies using RNA gel shift mobility assays showed that both HuR and CP1 can bind

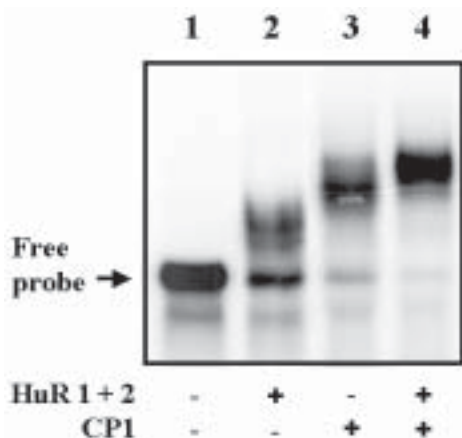


Fig. 2. Cooperative binding of HuR (domains 1 and 2) and CP1 to AR mRNA.

RNA gel shift mobility assay using ³²P-AR mRNA and purified recombinant HuR (domains 1 and 2) and CP1. The free probe is indicated (lane 1). It is partially shifted by either HuR 1+2 (lane 2) or CP1 (lane 3), with higher molecular weight complexes apparent in the gel. When both proteins are present the shifted band appears with increased intensity at a new position in the gel (lane 4), indicating simultaneous and cooperative binding.

simultaneously to the AR mRNA sequence, and also showed that the first two RNA-binding domains of HuR are sufficient for binding to the U-rich sequence. Furthermore, the interaction appears to be cooperative. Binding of either recombinant HuR or CP1 to the AR mRNA was augmented when added together in solution (Fig. 2).

These data pose interesting new questions from both a functional and structural perspective, and provide us with an excellent opportunity to examine the kinetics and complexities of this particular multiprotein-RNA interaction. How these proteins together act to regulate the AR mRNA stability and translational efficiency is the subject for future work in the Leedman laboratory. From a structural perspective it is interesting to contemplate the physical arrangement of HuR and CP1 and/or CP2 with the AR mRNA.

HuR and specific ARE binding interactions

Mammalian HuR is a 32-kDa protein ubiquitously expressed in all cell types (9). A number of mRNAs have been shown to be bound by HuR, including those encoding the β -adrenergic receptor, cyclins A, B1 and D1, *c-fos*, *c-myc*, neurofibromin and tumour necrosis factor (3). As well as upregulating early regulatory gene expression through binding to their mRNA instability elements (AREs), HuR is responsible for shuttling specific mRNAs between the nucleus and cytoplasm (11). There are four members of the Hu family of RNA-binding proteins. The others (HuB, HuC and HuD) are neuron-specific. They all contain three highly conserved RNA-binding domains belonging to the RNA recognition motif (RMM) family. The first two RMMs are linked by a short amino acid sequence (and cooperatively interact with RNA). Between the second and third RRM is a hinge region of 55 amino acid residues that contains signals for nuclear import and export (11).

The 3D structure of the first two binding domains of the closely related HuD protein bound to 11 nucleotide RNA fragments from the *c-fos* and tumour necrosis factor α (TNF α) AREs has been reported (12). This has provided insights into the mode of interaction of these domains with class I and class II AREs. To date, the third RRM of the Hu protein family has not been structurally characterised and the range of its RNA-binding specificity is not known. Although the third RRM domain has been proposed to bind both AREs and poly-A sequences (11,13), our work showed that it did not bind to AR mRNA on its own (14).

We are currently undertaking structural studies using NMR spectroscopy and X-ray crystallography to probe the specific binding interactions between HuR (and HuR domains), and the identified AR mRNA binding sequence. This sequence represents a class III type ARE, for which there is no Hu protein/RNA structure. We are also attempting to structurally analyse the third domain of HuR, which is the least characterised domain.

Poly-C-binding proteins (CPs)

The poly-C-binding proteins, CP1 and CP2, are highly conserved, widely distributed approximately 39 kDa RNA-binding proteins that bind to specific poly-C-containing regions of mRNA (15). The CPs bind to a variety of mRNAs, including tyrosine hydroxylase, erythropoietin, lipoxigenase, α -globin and several viral mRNAs, including

AR mRNA UC-rich sequence
(nt 3275-3325)

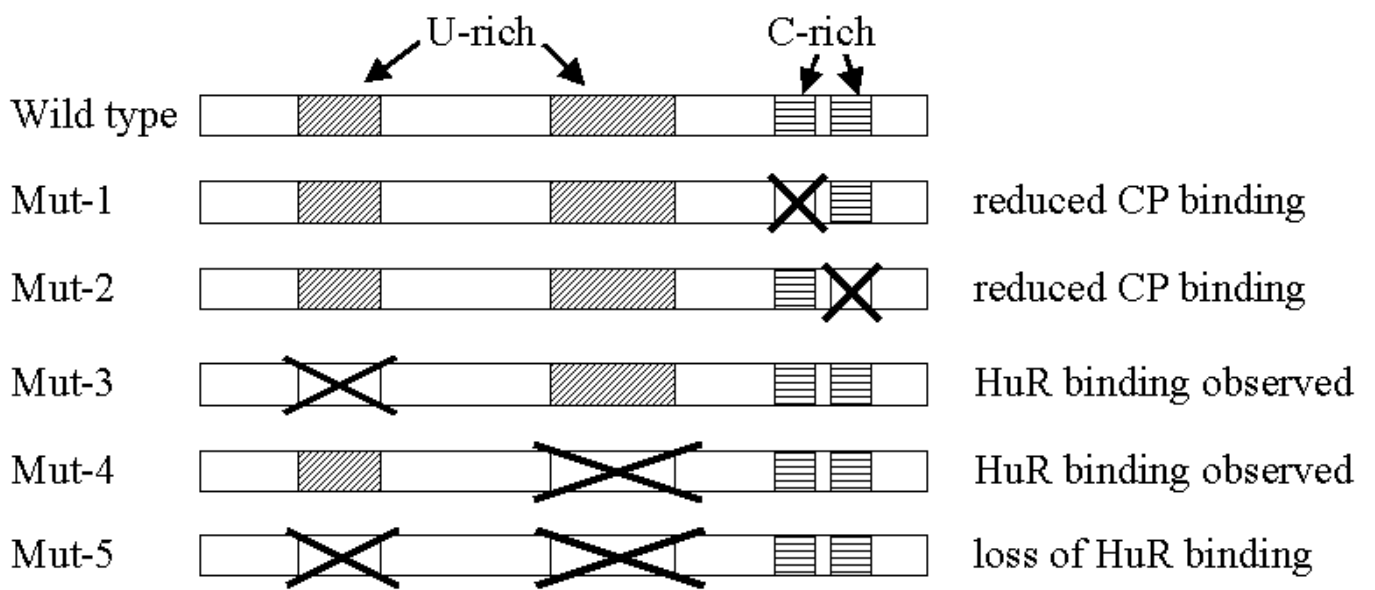


Fig. 1

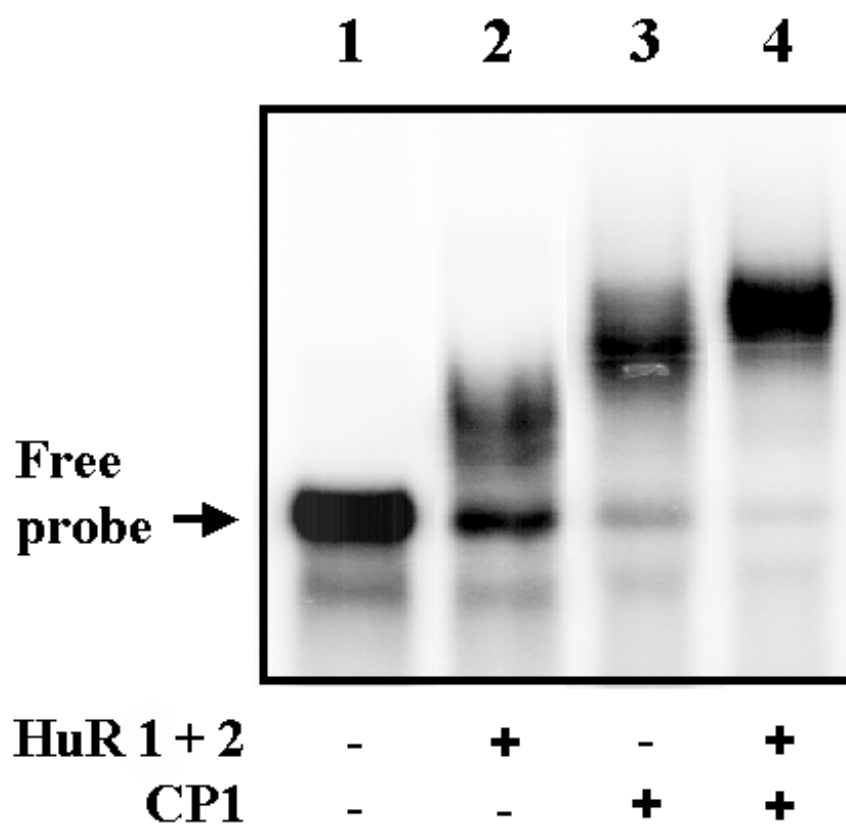


Fig. 2

