Eur Biophys J (2005) 34: 423–429
DOI 10.1007/s00249-005-0467-y

ARTICLE

M. Sidiqi · J. A. Wilce · C. J. Porter · A. Barker
P. J. Leedman · M. C. J. Wilce

Formation of an αCP1-KH3 complex with UC-rich RNA

Received: 18 November 2004 / Revised: 14 January 2005 / Accepted: 28 January 2005 / Published online: 9 March 2005

Abstract The αCP family of proteins [also known as poly(C)-binding or heterogeneous nuclear ribonucleoprotein E proteins] are involved in the regulation of messenger RNA (mRNA) stability and translational efficiency. They bind via their triple heterologous nuclear ribonucleoprotein K homology (KH) domain structures to C-rich mRNA, and are thought to interact with other mRNA-binding proteins as well as provide direct nuclease protection. In particular, αCP1 and αCP2 have been shown to bind to a specific region of androgen receptor (AR) mRNA, resulting in its increased stability. The roles of each of the KH motifs in the binding affinity and the specificity is not yet understood. We report the beginning of a systematic study of each of the αCP KH domains, with the cloning and expression of αCP1-KH2 and αCP1-KH3. We report the ability of αCP1-KH3, but not αCP1-KH2, to bind the target AR mRNA sequence using an RNA electrophoretic mobility gel shift assay. We also report the preparation of an αCP1-KH3/AR mRNA complex for structural studies. $^{1}H-^{15}N$ heteronuclear single quantum correlation NMR spectra of ^{15}N-labelled αCP1-KH3 verified the integrity and good solution behaviour of the purified domain. The titration of the 11-nucleotide RNA target sequence from AR mRNA resulted in a rearrangement of the $^{1}H-^{15}N$ correlations, demonstrating the complete binding of the protein to form a homogeneous protein/RNA complex suitable for future structural studies.

Keywords Poly(C)-binding protein · Heterogeneous nuclear ribonucleoprotein E1 · αCP1 · RNA binding · Heterologous nuclear ribonucleoprotein K homology domain

Abbreviations αCP: Poly(C)-binding protein · AR: Androgen receptor · DTT: Dithiothritol · EDTA: Ethylenediaminetetraacetic acid · GST: Glutathione S-transferase · hnRNP K: Heterologous nuclear ribonucleoprotein K · HSQC: Heteronuclear single quantum correlation · IPTG: Isopropyl-β-D-thiogalactoside · KH domain: hnRNP-K homology domain · mRNA: Messenger RNA · PAGE: Polyacrylamide gel electrophoresis · PBS: Phosphate-buffered saline · PMSF: Phenyl methyl sulfonyl fluoride · REMSA: RNA electrophoretic mobility shift assay · Tris: Tris(hydroxymethyl)aminomethane · UTR: Untranslated region

M. Sidiqi · J. A. Wilce · C. J. Porter · A. Barker
M. C. J. Wilce (✉)
School of Biomedical and Chemical Sciences,
The University of Western Australia, 35 Stirling Highway,
Crawley, Perth, WA, 6009, Australia
E-mail: mwilce@receptor.pharm.uwa.edu.au
Tel.: +61-8-93462981
Fax: +61-8-93463469

A. Barker · M. C. J. Wilce
School of Pharmacology and Medicine,
The University of Western Australia, 35 Stirling Highway,
Crawley, Perth, WA, 6009, Australia

P. J. Leedman
Laboratory for Cancer Medicine,
The UWA Centre for Medical Research,
The University of Western Australia, 35 Stirling Highway,
Crawley, Perth, WA, 6009, Australia

P. J. Leedman · M. C. J. Wilce
Western Australian Institute for Medical Research,
The University of Western Australia,
35 Stirling Highway, Crawley,
Perth, WA, 6009, Australia

Introduction

αCP1, also known as poly(C)-binding protein 1 and heterogeneous nuclear ribonucleoprotein E1, is a highly conserved and widely distributed approximately 42-kDa RNA-binding protein. αCP1 has been identified as a component of a ribonucleoprotein complex that assembles on poly(C) regions of several messenger RNAs (mRNAs), including tyrosine hydroxylase, erythropoietin, lipoxygenase, α-globin, β-globin, collagen α1(I), the

androgen receptor (AR) and several viral mRNAs resulting in alteration of mRNA stability or regulation of their translation rate (Paulding and Czyzyk-Krzeska 1999; Czyzyk-Krzeska and Bendixen 1999; Ostareck et al. 1997; Wang et al. 1995; Chkheidze et al. 1999; Yu and Russell 2001; Stefanovic et al. 1997; Yeap et al. 2002; Collier et al. 1998; Blyn and coworkers 1996, 1997; Gamarnik and Andino 1997; Graff et al. 1998; Spangberg and Schwartz 1999). This has thus been recognised as a very important point of regulation of gene expression (Guhaniyogi and Brewer 2001; Hollams et al. 2002). The formation of protein complexes on the mRNA in eukaryotic cells can make 1,000-fold differences in the half life of the mRNA and consequently enormous differences to the amount of protein product that is derived (Brennan and Steitz 2001).

As the name suggests, the αCP family (which also includes αCP2, αCP3, αCP4 and the earliest member to be characterised, the heterologous nuclear ribonucleoprotein K, hnRNP K) have a binding preference for poly(C) tracts of oligonucleotide, and bind both to RNA and to single-stranded DNA (Dejaard and Leffers 1996). They achieve this through their triplet K homology (KH) RNA-binding motif, as first identified in hnRNP K. The KH domains are 68–72 amino acid motifs (Musco et al. 1996; Adinolfi et al. 1999) and have been structurally characterised both in the absence (Musco and coworkers 1996, 1997; Lewis et al. 1999; Baber et al. 1999), and in the presence of RNA or DNA (Lewis et al. 2000; Liu et al. 2001; Braddock and coworkers 2002a, 2002b). This has revealed that the KH domain is a compact motif with a $\beta\alpha\alpha\beta\beta\alpha$ topology. A three-stranded antiparallel β-sheet is packed against three α-helices forming a hydrophobic groove where the oligonucleotide binds. The groove is bounded by two unstructured surface loops. The loop between α-helices 1 and 2 contains an invariant GXXG, crucial to oligonucleotide binding. That between β-strands 2 and 3 is of variable length in different KH domains and flanks the opposing edge of the RNA-binding groove (Lewis et al. 2000).

The relative contributions by the three KH domains in the αCP proteins to RNA-binding affinity and specificity is not yet clearly understood. Whilst all three of the three individual KH domains of hnRNP K have been shown to contribute significantly to poly(C) binding (Siomi et al. 1994), only the first and third KH domains of αCP1 and αCP2 have been shown to independently bind poly(C)-RNA with high affinity and specificity (Dejgaard and Leffers 1996). It may be, however, that the second KH domain also binds poly(C)-RNA when tethered by its neighbouring domains. The molecular basis for their poly(C) oligonucleotide specificity is not yet understood, but it is likely by analogy to homologous structures that they possess a narrow binding groove dictating a preference for pyrimidines and that specific hydrogen-bond interactions would complement cytosine hydrogen-bond donors and acceptors.

The interaction between αCP1 and a UC-rich region in AR mRNA has recently been shown to occur and potentially underlie AR mRNA stabilisation (Yeap et al. 2002). In addition, the αCP1 protein is likely to bind in close proximity to other proteins decorating the mRNA tract, including the nucleocytoplasmic shuttling protein HuR. A cooperative binding interaction has been observed in vitro for this pair of proteins in binding to the AR mRNA (nucleotides 3,275–3,325), and it is thus speculated that a unique tertiary complex could be forming that underlies AR mRNA stability (Wilce et al. 2002). Since the presence of the AR is integral to the progression of prostate cancer, such a complex would represent a novel drug target. Its disruption would be predicted to destabilise the AR mRNA and hence to reduce or obliterate the AR in prostate cancer cells, preventing the progression of the disease.

The current study describes the preparation of full-length αCP1 as well as its separate KH domains including αCP1-KH2 and αCP1-KH3. We investigated the ability of each of these three constructs to bind to the target AR mRNA sequence using an RNA electrophoretic mobility shift assay (REMSA). We also report the preparation of milligram quantities of an 11-nucleotide region of the target AR mRNA and its binding to αCP1-KH3 as monitored using NMR spectroscopy. This work represents the beginning of a thorough biophysical study of the interactions between each of the αCP1 KH domains and target RNA.

Experimental procedures

Protein expression and purification

Full-length αCP1 or individual KH domains were expressed as fusion proteins with glutathione S-transferase (GST). The DNA coding sequences comprising amino acids 1–356 (αCP1), 97–180 (αCP1-KH2), or 279–356 (αCP1-KH3) were cloned into pGEX-6P2 plasmids and expressed by the *E. coli* BL21 (Codon+) in Luria broth at 37°C (unlabelled protein). Protein expression was induced with 0.02 mM isopropyl-β-D-thiogalactoside (IPTG) at an optical density of 0.8 at 595 nm. The cells were harvested after 3 h of further growth by centrifugation and were resuspended in phosphate-buffered saline (PBS) (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO4, 1.8 mM KH_2PO4, pH 7.4) containing 0.5% Triton X-100 which was then lysed by French Pressing (SLM Instruments), supplemented with 0.5 mM phenyl methyl sulfonyl fluoride (PMSF) and, in the case of full-length αCP1 a cocktail of protease inhibitors was used, including aprotinin (2 μg/mL), leupeptin (2 μg/mL) and pepstatin (1 μg/mL). The supernatant was exposed to glutathione agarose beads equilibrated with PBS buffer at 4 °C overnight. The fusion protein was cleaved using 2 U Prescission protease (Amersham, Freiburg, Germany) in 50 mM tris(hydroxymethyl)aminomethane (Tris)–HCl, pH 7.0, 150 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA) and 1 mM dithiothritol

(DTT) at 4 °C overnight and was finally purified using a Sephadex 75 column (Pharmacia), or by anion-exchange chromatography in the case of full-length αCP1 (Mono-Q Pharmacia), and dialysed into phosphate buffer pH 6.0 (1 mM DTT, 25 mM NaH_2PO_4, 150 mM NaCl, 1 mM EDTA). The purified protein was concentrated with centrifugal concentrators of 3 K cut off (Millipore) and quantified using a detergent-compatible (BioRad, Mississauga, ON, Canada) protein assay.

Isotopically enriched ^{15}N-αCP1-KH3 was prepared for NMR spectroscopic studies, using cell culture grown in minimal media supplemented with vitamins and trace metals with $^{15}NH_4Cl$ as the sole nitrogen source. Cells were grown at 37 °C to an A_{595} nm of 7.7 in a fermentor (New Brunswick Scientific, NJ, USA). The temperature was changed to 30 °C before induction with 0.02 mM IPTG and the addition of 2 g $^{15}NH_4Cl$. Following induction, the growth was continued for 3 h to a final optical cell density of 13 and the cells were harvested using centrifugation following the same procedure as for the unlabelled protein except that the protein was finally purified using cation-exchange chromatography (Mono-S Pharmacia). The protein concentration was determined by measuring the absorbance at 280 nm using a molar extinction coefficient of 1,280 M^{-1} cm^{-1} estimated from the protein's tyrosine and tryptophan content (Gill and von Hippel 1989) and the protein was dialysed to a final NMR buffer of 90% sodium phosphate buffer (1 mM DTT, 10 mM NaH_2PO_4, 100 mM NaCl, 1 mM EDTA) and 10% D_2O (Isotec)

Preparation of RNA transcript for REMSA

A 51-nucleotide RNA corresponding to nucleotides 3,275–3,325 of the AR mRNA 3′ untranslated region (UTR) was generated by in vitro transcription of the relevant DNA cloned into pBLUESCRIPT II KS+ (Yeap et al. 2002). The pBLUESCRIPT vector alone was used as a control. Linearised plasmid DNA was transcribed in a 20-μl reaction volume with 20 U T7 RNA polymerase (Promega, Madison, WI, USA) at 37 °C for 60 min in the presence of 100 μCi (^{32}P UTP (Amersham Pharmacia Biotech., Chalfont, UK) and 2.5 mM each of ribosomal ATP, ribosomal CTP and ribosomal GTP (Amersham Pharmacia Biotech.), and 20 mM DTT. One unit of DNase 1 (RNase free) (Promega) was added for 10 min at 37 °C, followed by 5 min at 65°C. Loading dye [12 μl of 95% formamide, 20 mM EDTA, 0.3% bromophenol blue and xylene cyanol (wt/vol)] was added to the reaction mixture, which was heated to 80 °C for 3 min, before resolving labelled RNAs on 7 M urea/6% polyacrylamide gel, pre-electrophoresed in 1× 90 mM Tris borate/0.2 mM EDTA at 200 V for 20 min. Radiolabel transcripts were visualised using 2-min exposure of the gel to X-ray film. The full-length transcript was excised and eluted from the gel slice by shaking at 1,500 rpm for 4 h at 22 °C in sterile 0.5 M ammonium acetate and 1 mM EDTA. The RNA transcripts were recovered by ethanol precipitation, resulting in 1–4×10^{10} cpm/μg RNA (Thomson et al. 1999).

RNA electrophoretic mobility shift assay

The REMSA is a technique for investigating protein/RNA complexes. Purified full-length CP1, αCP1-KH2 or αCP1-KH3 was thawed on ice. A 1-μl aliquot of the protein (100 ng/μl) was transferred to a vial on ice containing cytoplasmic extraction buffer [10 mM *N*-(2-hydroxyethyl)piperazine-*N*′-ethanesulfonic acid pH 7.5, 3 mM $MgCl_2$, 14 mM KCl, 5% glycerol, 0.2% Nonidet P-40, 1 mM DTT; Thomson et al. 1999] and a cocktail of protease inhibitors (0.5 mM PMSF, 0.2 mM leupeptin, 0.2 mM aprotinin). A 1 μl aliquot of transfer RNA (2 μg/ml) was added to prevent nonspecific binding and the mixture was made up to a final volume of 10 μl with 1 μl of 10^4 cpm ^{32}P-labelled RNA. The mixtures were incubated on ice for 30 min and immediately after incubation 1 μl of loading dye was added and mixed gently. The reaction mixtures were loaded on a 6% nondenaturing acrylamide gel and run at 125 V in the cold room for 2 h. The gel was dried in a gel drier for 20 min at 80 °C, and was then exposed to a phosphorimager plate overnight for detection of protein and RNA complexes.

Preparation of 11-nucleotide αCP1 target site from AR mRNA

The RNA oligonucleotide representing the 11 last nucleotides of the AR mRNA region shown to bind αCP1 (nucleotides 3,315–3,325, 5′-UUCCCUCCCUA-3′) was purchased from Dharmacon in crude form and further purified by denaturing polyacrylamide gel electrophoresis (PAGE). After the separation of the sample by 20% PAGE, the band was visualised by UV shadowing once the RNA had run approximately one half way down the gel. The sample was recovered by excising the appropriate band, which was then crushed and eluted overnight in 0.3 M sterile sodium acetate. The eluent was filtered and desalted using a reverse-phase solid-extraction cartridge (C18 Sepak cartridge, Waters). The eluted fractions were lyophilised and dissolved in distilled water for quantification using UV spectroscopy. RNA concentrations were determined by measuring the absorbance at 260 nm and assuming one absorbance unit to be equivalent to 34 μg/ml.

Preparation of αCP1-KH3 and αCP1-KH3/RNA complexes for NMR

The ^{15}N-labelled αCP1-KH3 was prepared at 200–500 μM protein in 10 mM sodium phosphate buffer pH 6.0, containing 100 mM NaCl, 2 mM EDTA, 2 mM

DTT and 10% D_2O. This sample was used to collect spectra for $^{1}H-^{15}N$ heteronuclear single quantum coherence (HSQC) spectra for *apo*-αCP1-KH3.

Aliquots of RNA were added in a stepwise manner to the ^{15}N-αCP1-KH3. Desired ratios of the protein and the RNA were combined in volumes of 3–4 ml such that upon concentration to 500 µl, the protein concentration was 300 µM and the RNA concentration was 75 µM. The final protein to RNA ratio was approximately 1:1. The solutions were concentrated using 1-kDa molecular weight cut-off concentrators to a final volume of 500 µl. To all samples, D_2O was added to 10%. The titrations were monitored by collecting $^{1}H-^{15}N$ HSQC spectra at each titration point.

NMR spectroscopy

Two-dimensional $^{1}H-^{15}N$ HSQC spectra were acquired with a Bruker Avance600 NMR spectrometer operating at 25 °C, using adapted versions of the published pulse sequences (Bax et al. 1990; Norwood et al. 1990). Water suppression was achieved by replacing the final 90° pulse with a Watergate sequence (Piotto et al. 1992). Relaxation effects were minimised by setting the evolution period for $^{15}N-^{1}H$ one-bond couplings to 2.3 ms, slightly shorter than $1/4\ J_{NH}$.The ^{15}N-decoupling during acquisition was achieved using the GARP decoupling scheme (Shaka et al. 1985). The ^{1}H carrier frequency was set to that of the water resonance and the ^{15}N carrier frequency was set at a frequency between the Arg Nϵ and backbone amide resonances. Time-proportional phase incrementation (20) was used to achieve quadrature detection over spectral widths of 12.25 ppm (7.3 kHz) for F_2 and 40.0 ppm (2.4 kHz) for F_1. A total of 32 scans per increment were collected over 256 t_1 increments of 2,048 complex data points. Spectral processing was carried out using XWINNMR software. Strip transformation of the regions between 6.3 and 9.8 (F_2) and 103.5 and 132.0 ppm (F_1) and zero filling into 2,048×2,048 real data points was used to increase digital resolution, while spectral resolution was enhanced by apodisation with a Lorentz–Gauss function in both dimensions prior to Fourier transformation. Baselines were corrected after phase correction by subtracting a third-order polynomial fitted to the baseline.

Results and discussion

Preparation of protein and RNA

Full-length αCP1 protein readily expressed as a GST fusion protein, which was partially soluble in the presence of the detergent Triton X-100. In the absence of the detergent all expressed protein was present in the insoluble fraction of the cell lysate and as a result detergent was also included in the application of the cell lysate to the GSH resin. In addition, the purified product showed susceptibility to degradation despite the inclusion of protease inhibitors. The sodium dodecyl sulfate PAGE revealed the full-length protein and a proteolytic product, which was identified using N-terminal sequencing as an N-terminal fragment of the protein. For the mobility shift assay the amount of the proteolytic product present was negligible (Fig. 1a).

Preparation of αCP1-KH2 involved using the same procedure as for the full-length protein except that the protein being approximately 8 kDa was purified from the 27-kDa GST contaminant using gel filtration chromatography. Expressed αCP1-KH2 was present in both the soluble and the insoluble fraction of the cell lysate, resulting in a yield of about 1 mg of highly pure protein per litre of culture (Fig. 1b).

αCP1-KH3 overexpressed well in both rich media and minimal media; however, it too was present in both the soluble and the insoluble fraction of the cell lysate in both media even with the induction of the protein at a lower temperature. The total protein yield was 2 mg of purified protein per litre of culture (Fig. 1b). It was

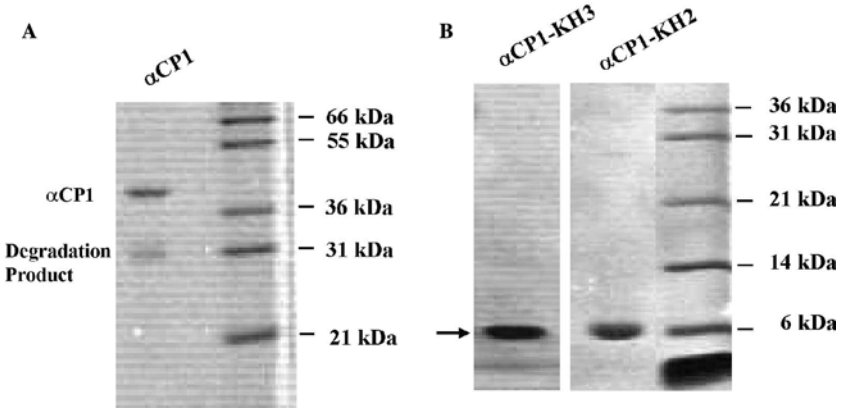

Fig. 1 Purification of αCP1, αCP1-KH3 and αCP1-KH2 using the pGex6P2 system. **a** The sodium dodecyl sulfate polyacrylamide gel electrophoresis gel shows αCP1 at the expected size of 42 kDa and the degradation product at just below 31 kDa after gel filtration. **b** The positions of pure αCP1-KH3 and αCP1-KH2 are indicated by the *arrow* at the expected molecular weight of approximately 8 kDa. Molecular weight markers are shown on the *right*. See the text for an explanation of the terms

hypothesised that cell growth in minimal media and the induction of the protein at 30°C would enhance the solubility of the expressing protein, but in our experiments there were no major improvements in the solubility level. For the purpose of the RNA-binding studies we prepared stocks of 0.6 mM unlabelled protein and 0.4 mM ^{15}N-labelled protein.

RNA-binding studies

The REMSA was used to examine the ability of full length αCP1-KH2 and αCP1-KH3 to bind to a 51-nucleotide UC-rich sequence from the 3′UTR of AR mRNA (nucleotides 3,275–3,325). The binding by full-length αCP1 has previously been demonstrated, but whether the separate KH domains bind to this sequence had not yet been tested (Yeap et al. 2002). The results are shown in Fig. 2. As expected, binding to the target RNA by full-length αCP1 is indicated by a substantial and quantitative shift in its mobility (lane 4). A quantitative shift of this RNA target by αCP1-KH3 is also clearly discernible (lane 3), although the degree of change in its mobility is not as marked as that seen with full-length αCP1. This difference in the relative mobility shift is due to the greater size of the full-length protein (37.5 kDa) in comparison with αCP1-KH3 (8 kDa). On the other hand, αCP1-KH2 exhibits no binding (lane 2), even though the protein is present in excess over RNA. Neither full-length αCP1 nor αCP1-KH3 showed any binding interaction to pBLUESCRIPT RNA alone (results not shown), demonstrating that the binding interaction occurs with the target RNA.

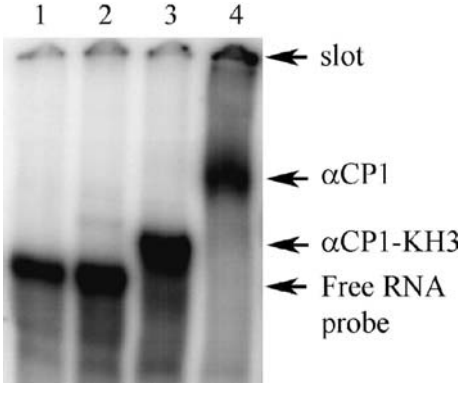

Fig. 2 Binding studies of αCP1 and individual KH domains to the 3′UTR region of androgen receptor messenge RNA. A typical RNA electrophoretic mobility shift assay gel is shown, in which binding by αCP1 or isolated domains of αCP1 to radioactively labelled RNA (with the sequence 5′-CUGGGUUUUUUUUUC-UCUUUCUCUCCUUUCUUUUUCUUCUUCCCUCCCUA-3′ appended to the pBLUESCRIPT vector) in the presence of excess transfer RNA as a nonspecific competitor were examined. *Lane 1*, probe only; *lane 2*, also contains 100 ng αCP1-KH2; *lane 3*, also contains 100 ng αCP1-KH3; *lane 4*, also contains 100 ng αCP1 (full length)

This extends the finding of Dejgaard and Leffers (1996), who observed at best weak binding to poly(C)-RNA by isolated αCP1-KH2 using a dot-blot assay. The absence of binding by αCP1-KH2 also indicates that the binding observed by the other species is not due to nonspecific protein/RNA interactions in the buffer conditions used here. Thus, αCP1-KH3 is capable of binding independently and specifically to sequences within the AR mRNA 3′UTR that are also contacted by full-length αCP1.

Formation of an αCP1-KH3/11-nucleotide RNA complex

Mutational analysis of the UC-rich 51-nucleotide sequence in the 3′UTR of AR mRNA has previously shown that the binding of αCP1 or αCP2 to this region is dependent on the presence of the two cytosine triplets at the end of the sequence (Yeap et al. 2002). The 11 last nucleotides (5′-UUCCCUCCCUA-3′) were therefore prepared as the target for αCP1-KH3 binding. The ^{15}N-labelled αCP1-KH3 could then be monitored using NMR spectroscopy to observe the effects of the addition of the oligonucleotide to the sample.

The ^{15}N-labelled αCP1-KH3 gave rise to well-resolved HSQC spectra (Fig. 3a), showing excellent dispersion in both ^{1}H and ^{15}N dimensions as well as narrow line widths. The dispersion, particularly in the ^{1}H dimension, indicates that the domain is likely to be folded in its correct secondary and tertiary structure and the narrow lines are consistent with this construct behaving as a monomer in solution. Seventy-nine single ^{1}H–^{15}N amide crosspeaks are observed, as would be expected for this 83-residue construct containing three prolines. These and the N-terminal amine do not give rise to an amide crosspeak. In addition, 12 doublet ^{1}H–^{15}N amine crosspeaks are distinguishable (between 108 and 114 ppm in the ^{15}N dimension), which represent signals from the six glutamine and six asparagine side chain amines.

Upon the addition of RNA to the ^{15}N-labelled αCP1-KH3 sample, the positions of the crosspeaks changed, reflecting altered electronic environments for many backbone NH groups in the protein. The final titration point is shown in Fig. 3b. The ^{1}H−^{15}N HSQC spectrum remains well dispersed and has 79 well + resolved crosspeaks. However at least 35 of the crosspeaks representing backbone NH correlations have changed position. This demonstrates that the protein is fully complexed with the RNA (i.e. no evidence of heterogeneity) and that the complex retains good solution characteristics with no evidence of aggregation or the formation of larger complexes. In addition, the crosspeak movement shows that almost half the backbone NH residues experience an altered electronic environment upon interaction with RNA—more than would be directly at the protein/RNA interface. This is unsurprising considering the long-range electrostatic effects

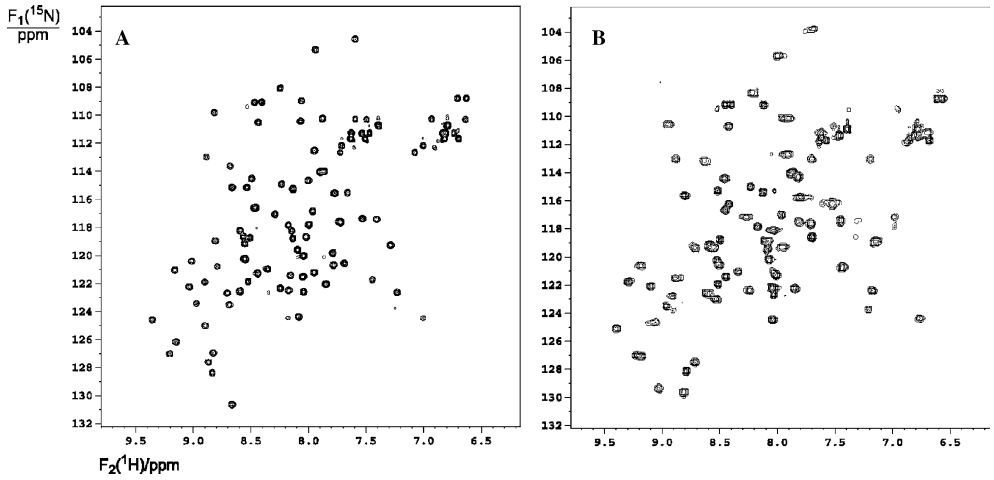

Fig. 3 The $^1H-^{15}N$ heteronuclear single quantum correlation spectra recorded at 25 °C for ^{15}N-labelled αCP1-KH3 before and after the addition of the 11-nucleotide RNA of sequence 5′-UUCCCUCCCUA-3′. **a** The uncomplexed spectrum and **b** the final titration point with αCP1-KH3 fully complexed with RNA. The crosspeaks on both spectra account for all of the expected resonances in the protein, possess narrow linewidths and are well dispersed. The movement of almost half of the peaks upon complex formation with RNA is consistent with a tight protein/RNA binding interaction

that would be expected to arise from the oligonucleotide's phosphate backbone. Spectra acquired at intermediate stages of the titration showed no evidence of a gradual movement of the crosspeaks from their starting to finishing positions. This would be typical of a weak interaction in which the chemical shift values represent averaged positions in this fast-exchange regime. Rather, the peaks disappeared and reappeared in new positions, suggesting a tight binding interaction and slow exchange relative to the NMR timescale.

Conclusions

This study represents the beginning of a functional and structural examination of all three KH domains of each αCP isoform. The basis for the RNA-binding affinity and specificity of the three KH domains will allow us to predict the occurrence of αCP interactions with mRNA and begin to rationalise its role in mRNA stability and translational efficiency. We have shown that like full-length αCP1, αCP1-KH3 binds to UC-rich RNA representing nucleotides 3,275–3,325 of AR mRNA. In contrast, αCP1-KH2 shows no evidence of binding, suggesting either a marked difference in oligonucleotide specificity or a non-RNA binding role in the full-length αCP1 protein.

NMR spectroscopy was also used to confirm the ability of αCP1-KH3 to bind to an AR mRNA sequence, as well as to demonstrate its complete complex formation with the 11-nucleotide sequence at the final titration point. It also showed that almost half of the αCP1-KH3 backbone NH resonances are affected by RNA binding. Assignment of the crosspeaks would be required to determine which residues these are. The current studies show that this protein and this protein/RNA complex would be highly amenable to further NMR studies. The complex is thus also ideal for crystallographic trials for its structure determination using X-ray crystallography. We have previously found that NMR evidence for protein/oligonucleotide complex formation, solubility and homogeneity is an excellent starting point for crystallographic studies (Vivian et al. 2001).

Structural studies of αCP1-KH3/AR mRNA will be a step towards describing the multiprotein AR mRNA complex that influences its stability and possibly its translational efficiency in vivo. Structural insight into such a complex may pave the way for the development of novel therapeutics aimed at disrupting the complex. This could lead to AR mRNA instability and reduce the amount of the AR in prostate cancer cells.

Acknowledgements This work was funded by an Australian Research Council Grant (M.C.J.W., J.A.W. and P.J.L.), an Australian Research Council Fellowship (J.A.W.), an Australian Postgraduate Award (M.S.) and a Small Grant awarded by the University of Western Australia (J.A.W.).

References

Adinolfi S, Bagni C, Castiglione Morelli MA, Fraternali F, Musco G, Pastore A (1999) Novel RNA-binding motif: the KH module. Biopolymers 51:153–164

Baber JL, Libutti D, Levens D, Tjandra N (1999) High precision solution structure of the C-terminal KH domain of heterogeneous nuclear ribonucleoprotein K, a c-myc transcription factor. J Mol Biol 289:949–962

Bax A, Ikura M, Kay LE, Torchia DA, Tschudin R (1990) Comparison of different modes of two-dimensional reverse correlation NMR for the study of proteins. J Magn Reson 86:304–318

Blyn LB, Swiderek KM, Richards O, Stahl DC, Semier BL, Ehrenfeld E (1996) Poly(rC) binding protein 2 binds to stem-loop IV of the poliovirus RNA 5′ noncoding region: identification by automated liquid chromatography-tandem mass spectrometry. Proc Natl Acad Sci USA 93:11115–11120

Blyn LB, Towner JS, Semier BL, Ehrenfeld E (1997) Requirement of poly(rC) binding protein 2 for translation of poliovirus RNA. J Virol 71:6243–6246

Braddock DT, Baber JL, Levens D, Clore GM (2002a) Molecular basis of sequence-specific single-stranded DNA recognition by KH domains: solution structure of a complex between hnRNP K KH3 and single-stranded DNA. EMBO J 21:3476–3485

Braddock DT, Louis JM, Baber JL, Levens D, Clore GM (2002b) Structure and dynamics of KH domains from FBP bound to single-stranded DNA. Nature 415:1051–1056

Brennan CM, Steitz JA (2001) HuR and mRNA stability. Cell Mol Life Sci 58:266–277

Chkheidze AN, Lyakhov DL, Makeyev AV, Morales J, Kong J, Liebhaber SA (1999) Assembly of the a-globin mRNA stability complex reflects binary interaction between the pyrimidine-rich 3-untranslated region determinant and poly©) binding protein αCP. Mol Cell Biol 19:4572–4581

Collier B, Goobar-Larsson L, Sokolowski M, Schwartz S (1998) Translational inhibition in vitro of human papillomavirus type 16 L2 mRNA mediated through interaction with heterogenous ribonucleoprotein K and poly(rC)-binding proteins 1 and 2. J Biol Chem 273:22648–22656

Czyzyk-Krzeska MF, Bendixen AC (1999) Identification of the poly©) binding protein in the complex associated with the 3′-untranslated region of erythropoietin mRNA. Blood 93:2111–2120

Dejgaard K, Leffers H (1996) Characterisation of the nucleic-acid-binding activity of KH domains. Different properties of different domains. Eur J Biochem 241:425–431

Gamarnik AV, Andino R (1997) Two functional complexes formed by KH domain containing proteins with the 5' noncoding region of poliovirus RNA. RNA 3:882–892

Gill SC, von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem 182:319–326

Graff J, Cha J, Blyn LB, Ehrenfeld E (1998) Interaction of poly(rC) binding protein 2 with the 5′ noncoding region of hepatitis A virus RNA and its effects on translation. J Virol 72:9668–9675

Guhaniyogi J, Brewer G (2001) Regulation of mRNA stability in mammalian cells. Gene 265:11–23

Hollams EM, Giles KM, Thomson AM, Leedman PJ (2002) mRNA stability and the control of gene expression: implications for human disease. J Neurochem Res 27:957–980

Lewis HA, Chen H, Edo C, Buckanovich RJ, Yang YY, Musunuru K, Zhong R, Darnell RB, Burley SK (1999) Crystal structures of Nova-1 and Nova-2 K-homology RNA-binding domains. Struct Fold Des 15:191–203

Lewis HA, Musunuru K, Jensen KB, Edo C, Chen H, Darnell RB, Burley SK (2000) Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome. Cell 100:323–332

Liu Z, Luyten I, Bottomley MJ, Messias AC, Houngninou-Molango S, Sprangers R, Zanier K, Kramer A, Sattler M (2001) Structural basis for recognition of the intron branch site RNA by splicing factor 1. Science 294:1098–1102

Musco G, Stier G, Joseph C, Castiglione Morelli MA, Nilges M, Gibson TJ, Pastore A (1996) Three-dimensional structure and stability of the KH domain: molecular insights into the fragile X syndrome. Cell 85:237–245

Musco G, Kharrat A, Stier G, Fraternali F, Gibson J, Nilges M, Pastore A (1997) The solution structure of the first KH domain of FMR1, the protein responsible for the fragile X syndrome. Nat Struct Biol 4:712–716 Erratum in: Nat Struct Biol 1997, 4:840

Norwood TJ, Boyd J, Heritage JE, Soffe N, Campbell ID (1990) Comparison of techniques for ^{1}H-detected heteronuclear $^{1}H-^{15}N$ spectroscopy. J Magn Reson 87:488–501

Ostareck DH, Ostareck-Lederer A, Wilm M, Theile, BJ, Mann M, Hentze MW (1997) mRNA silencing in erythroid differentiation: hnRNP K and hnRNP E1 regulate 15-lipoxygenase translation from the 3′ end. Cell 89:597–606

Paulding WR, Czyzyk-Krzeska MF (1999) Regulation of tyrosine hydroxylase mRNA stability by the protein-binding, pyrimidine-rich sequence in the 3′-untranslated region. J Biol Chem 274:2532–2538

Piotto M, Saudek V, Sklenar V (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. J Biomol NMR 2:661–665

Shaka AJ, Barker PB, Freeman R (1985) Computer-optimised decoupling scheme for wideband applications and low-level operation. J Magn Reson 64:547–552

Siomi H, Choi M, Siomi MC, Nussbaum RL, Dreyfuss G (1994) Essential role for KH domains in RNA binding: impaired RNA binding by a mutation in the KH domain in FMR1 that causes fragile X syndrome. Cell 77:33–39

Spangberg K, Schwartz S (1999) Poly(C)-binding protein interacts with the hepatitis C virus 5' untranslated region. J Gen Virol 80:1371–1376

Stefanovic B, Hellerbrand C, Holcik M, Briendl M, Liebhaber SA, Brenner DA (1997) Posttranscriptional regulation of collagen alpha1 (I) mRNA in hepatic stellate cells. Mol Cell Biol 17:5201–5209

Thomson AM, Rogers JT, Walker CE, Staton JM, Leedman PJ (1999) Optimised RNA gel-shift and UV cross-linking assays for characterization of cytoplasmic RNA-protein interactions. Biotechniques 27:1032–1040

Vivian JP, Wilce JA, Hastings AF, Wilce MCJ (2001) Crystallization of the *B. subtilis* RTP/DNA complex. Acta Crystallogr D 57:421–424

Wang X, Kiledjian M, Weiss IM, Liebhaber SA (1995) Detection and characterization of a 3′-untranslated region ribonucleoprotein complex associated with human α-globin mRNA stability. Mol Cell Biol 15:1769–1777

Wilce JA, Leedman PJ, and Wilce MCJ (2002) RNA-binding proteins that target the androgen receptor mRNA. IUBMB Life 54:345–349

Yeap BB, Voon DC, Vivian JP, McCulloch RK, Thomson AM, Giles KM, Czyzyk-Krzeska MF, Furneaux H, Wilce MC, Wilce JA, Leedman PJ. (2002) Novel binding of HuR and poly(C)-binding protein to a conserved UC-rich motif within the 3′-untranslated region of the androgen receptor messenger RNA. J Biol Chem 277:27183–27192

Yu J, Russell JE (2001) Structural and functional analysis of an mrnp complex that mediates the high stability of human beta-globin mRNA. Mol Cell Biol 21:5879–5888