

Polyadenylation proteins CstF-64 and τ CstF-64 exhibit differential binding affinities for RNA polymers

Roberto R. MONAREZ, Clinton C. MacDONALD and Brinda DASS¹

Department of Cell Biology & Biochemistry, Texas Tech University Health Sciences Center, 3601 4th Street, Lubbock, TX 79430, U.S.A.

CstF-64 (cleavage stimulation factor-64), a major regulatory protein of polyadenylation, is absent during male meiosis. Therefore a paralogous variant, τ CstF-64 is expressed in male germ cells to maintain normal spermatogenesis. Based on sequence differences between τ CstF-64 and CstF-64, and on the high incidence of alternative polyadenylation in testes, we hypothesized that the RBDs (RNA-binding domains) of τ CstF-64 and CstF-64 have different affinities for RNA elements. We quantified K_d values of CstF-64 and τ CstF-64 RBDs for various ribopolymers using an RNA cross-linking assay. The two RBDs had similar affinities for poly(G)₁₈, poly(A)₁₈ or poly(C)₁₈, with affinity for poly(C)₁₈ being the lowest. However, CstF-64 had a higher affinity for poly(U)₁₈ than τ CstF-64, whereas it had a lower affinity for

poly(GU)₉. Changing Pro-41 to a serine residue in the CstF-64 RBD did not affect its affinity for poly(U)₁₈, but changes in amino acids downstream of the C-terminal α -helical region decreased affinity towards poly(U)₁₈. Thus we show that the two CstF-64 paralogues differ in their affinities for specific RNA sequences, and that the region C-terminal to the RBD is important in RNA sequence recognition. This supports the hypothesis that τ CstF-64 promotes germ-cell-specific patterns of polyadenylation by binding to different downstream sequence elements.

Key words: cleavage stimulation factor (CstF), male germ cell, mRNA processing, polyadenylation, RNA-binding domain (RBD), spermatogenesis.

INTRODUCTION

Polyadenylation is the process by which approx. 200 adenylate residues are added to 3'-ends of most eukaryotic pre-mRNAs [1,2]. This process is crucial for proper transcription, splicing, transport, translation and stability of mRNAs [3–12]. In mammals, polyadenylation requires at least five distinct components: the CPSF (cleavage and polyadenylation specificity factor), the CstF (cleavage stimulation factor), CF I_m and CF II_m (mammalian cleavage factors I and II) and the PAP [poly(A) polymerase] [1,2]. Other proteins may also contribute to the efficiency or specificity of polyadenylation [13–19].

Two proteins of the polyadenylation machinery are known to bind to important RNA sequence elements near polyadenylation sites. The 160 kDa subunit of CPSF binds to the canonical AAUAAA element found usually 12–30 nt upstream of the cleavage site of a pre-mRNA [20,21], while CstF binds to a U- or G/U-rich DSE (downstream sequence element) through its 64 kDa subunit, CstF-64 [22–24]. CstF-64 contains an N-terminal RBD (RNA-binding domain) consisting of an RRM (RNA recognition motif) of the type found in the U1A small ribonucleoprotein [23,25–28], a CstF-77-binding domain [29], a repeated MEARA helical region [23,30] and a conserved C-terminal domain reported to bind the transcription factor PC-4 [31].

Meiotic and post-meiotic male germ cell pre-mRNAs have a higher incidence of non-AAUAAA polyadenylation signals than somatic cells [32,33]. There is also a much higher incidence of alternative polyadenylation in male germ cells [34,35]. We described earlier a germ-cell-expressed paralogue of CstF-64 that we designated τ CstF-64 [36,37], which might contribute to polyadenylation of non-AAUAAA pre-mRNAs [32]. X-linked genes including CstF-64 are not expressed in meiotically dividing

male germ cells due to the unique process of male X-inactivation [38–40]. To compensate for X chromosomal inactivation, many X-linked genes have autosomal paralogues that are expressed in male germ cells [41]. Similarly, τ CstF-64 is on an autosome (*Cstf2t* in mice, *CSTF2T* in humans) and has diverged from the somatic CstF-64 [37,42]. Hence, we have hypothesized that τ CstF-64 might be at least partially responsible for alternative polyadenylation and polyadenylation of non-AAUAAA pre-mRNAs in male germ cells [32].

The cloning and sequencing of the mouse and human τ CstF-64 cDNAs showed that the translated sequence was similar to other known isoforms of CstF-64 (72.9 and 74% identical with mouse and human respectively, [37,42]). Alignment of CstF-64 and τ CstF-64 showed that they contained common features such as similar RBDs, CstF-77 interaction domains, MEARA repeat regions, and C-terminal domains. The RRM-type RBDs of CstF-64 and τ CstF-64 consist of two α -helices (α_1 and α_2 ; Figure 1) and four β -sheets (β_1 – β_4 ; Figure 1). α -Helices α_1 and α_2 stabilize the RNA-binding β -sheet surface from behind (Figure 1B, [28]). Recently, Pérez Cañadillas and Varani [27] determined the solution structure of the CstF-64 RBD and found that it contained a third α -helix (α_3) that is C-terminal to the core RBD. Those authors proposed a model in which α_3 lay across the RNA-contact β -sheet surface of the CstF-64 RBD and precluded binding of RNA until moved by some signal from elsewhere in the molecule.

Alignment of CstF-64 and τ CstF-64 RBD sequences suggests that they have nearly identical three-dimensional structures. However, the core RBD of τ CstF-64 differs from that of CstF-64 in that it has a serine residue at amino acid 41 rather than a proline (Figure 1A, [37,42]). The third α -helices (α_3) in τ CstF-64 and CstF-64 are the same, but there are a number of amino acid

Abbreviations used: CF I_m, mammalian cleavage factor I; CPSF, cleavage and polyadenylation specificity factor; CstF, cleavage stimulation factor; DSE, downstream sequence element; GST, glutathione S-transferase; RBD, RNA-binding domain; RRM, RNA recognition motif; SELEX, systematic evolution of ligands by exponential enrichment; SV40, simian virus 40.

¹ To whom correspondence should be addressed (email brinda.dass@ttuhsc.edu).

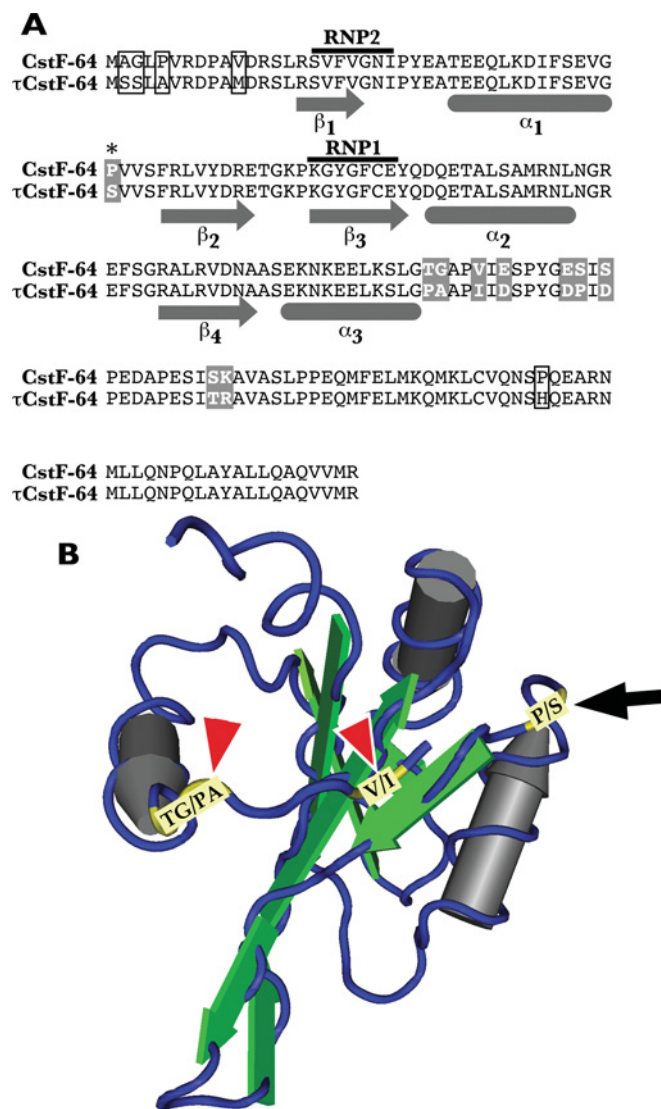


Figure 1 Amino acid differences of the RBD between the CstF-64 and τ CstF-64 isoforms

(A) CstF-64 and τ CstF-64 amino acid sequences. Constructs used in the present study end at amino acid 180 as shown. α -Helical and β -sheet regions are indicated; RNP1 and RNP2 indicate conserved sequence motifs in RRM-type RNA-binding proteins. Amino acids that differ between CstF-64 and τ CstF-64 are boxed or boxed in grey; amino acids that were altered in the present study are boxed in grey, and amino acid 41 is indicated by an asterisk. (B) Three-dimensional representation of the CstF-64 RBD [27] highlighting the differences between CstF-64 and τ CstF-64. α -Helices are grey cylinders, β -sheets are green ribbons, and the peptide backbone is a blue thread. Amino acids that differ are yellow. The black arrow indicates amino acid 41 (P/S), and amino acids 106, 107 and 110 (TG/PA and V/I) are indicated by red arrowheads.

differences between CstF-64 and τ CstF-64 immediately downstream of α_3 (Figure 1A). These variations between τ CstF-64 and CstF-64 suggest a potential for the two proteins to have differing affinities for specific RNA sequences.

Currently, no consensus 'testis-specific polyadenylation signal' has been determined [32,33]. This led us to consider an alternative approach for determining potential sequences based on the binding preferences of τ CstF-64. In the present study, we examined the binding affinities of several ribopolymers to CstF-64 and τ CstF-64. A UV cross-linking competition assay was used to determine binding of the CstF-64 and τ CstF-64 RBDs to various RNA substrates. We found that CstF-64 and τ CstF-64 had nearly

identical affinities for each of the following substrates, poly(A), poly(G) and poly(C), with the affinity for poly(C) being significantly lower than for any of the other elements. However, we found that CstF-64 bound poly(U) more efficiently than poly(GU). On the other hand, it was the reverse for τ CstF-64: poly(GU) was bound more readily than poly(U). When using mutants to examine specific amino acids that might contribute to such specificities, we saw no change in binding of CstF-64 to poly(U) by changing the Pro-41 of CstF-64 to a serine residue. However, changing nine amino acids in the region adjacent to α_3 decreased slightly the affinity of CstF-64 for poly(U), suggesting that the region in or around α_3 contributed to the different RNA-binding specificities of CstF-64 and τ CstF-64. The different binding affinities of CstF-64 and τ CstF-64 for different RNA elements are consistent with the hypothesis that τ CstF-64 promotes germ cell-specific patterns of polyadenylation by binding to different sequence elements downstream of polyadenylation sites than does CstF-64.

EXPERIMENTAL

Recombinant protein purification

RBDs of mouse CstF-64 and mouse τ CstF-64 were prepared as GST (glutathione S-transferase) fusion proteins. PCR was performed on previously described clones of mouse CstF-64 and mouse τ CstF-64 cDNAs [37,42] with primers that inserted a BamHI site upstream of the ATG and an EcoRI site 550 bp downstream of the region that contained the RBDs (CstF-64: Primer 1, 5'-CCTCGGGGGGATCCATGGCGGGTTTG-3' and Primer 2, 5'-CAGCGCAATCTCAGAAATTCACAATTCTC-3'; τ CstF-64: Primer 3, 5'-GTCCTCTCGGATCCATCATGTGCGAGTTTG-3' and Primer 4, 5'-GCAATCTCTGAATTCATGATTCTCATC-3'). These PCRs generated fragments of 579 bp and 580 bp for mouse CstF-64 RBD and τ CstF-64 RBD respectively, which were cloned into pCR-TOPO II vector using the TOPO-TA cloning system (Invitrogen, Carlsbad, CA, U.S.A.). These constructs were used as the source to clone CstF-64 and τ CstF-64 RBDs into the BamHI and EcoRI sites of pGEX-2TK expression vector (Amersham Biosciences, Piscataway, NJ, U.S.A.) ensuring they were in-frame with the GST tag. Mouse CstF-64 RBD and τ CstF-64 RBD in pGEX-2TK plasmid DNAs were electroporated using a Cell-Porator[®] *Escherichia coli* pulser (Life Technologies, Gaithersburg, MD, U.S.A.) into BL21 (DE3) Star One Shot cells (Invitrogen). These cells were induced at mid-exponential phase by the addition of 0.25 mM (final concentration) IPTG (isopropyl β -D-thiogalactoside) at 37 °C for 5 h. Bacterial pellets were resuspended in lysis buffer (50 mM Tris/HCl, pH 7.5, 0.1 M NaCl, 1.0 mM EDTA, 0.5% Nonidet P40, 1 mM *o*-phenanthroline, 50 μ M aprotinin and 1 mM EGTA), sonicated on ice, and cleared by centrifugation (158 000 g_{av} , 45 min and 4 °C; Beckman Ti45 rotor; 45 000 rev./min). GST-tagged recombinant proteins were isolated in batch using the glutathione-Sepharose 4B system (Amersham Biosciences) according to the manufacturer's instructions. Protein concentration was estimated using the Bio-Rad Protein Assay reagent (Bio-Rad Laboratories, Hercules, CA, U.S.A.) and by comparison with known amounts of BSA isolated by SDS/12.5% PAGE and visualized using Coomassie Brilliant Blue R-250 (Fisher Scientific, Hampton, NH, U.S.A.).

RNA ligands

The four different RNA ribopolymers, poly(U)₁₈, poly(A)₁₈, poly(G)₁₈ and poly(GU)₉ (Table 1) were synthesized using the T7-MEGAshortscript High Yield Transcription kit (Ambion, Austin, TX, U.S.A.) as directed. Poly(C)₁₈ was synthesized

Table 1 Oligonucleotides used in RNA ligand preparation using the T7-MEGAShortscript High Yield Transcription kit

Strand	Sequence
Top	5'-TAATACGACTCACTATAGGG-3'
Bottom	
Poly(U) ₁₈	5'-AAAAAAAAAAAAAAAAACCCCTATAGTGAGTCGTATTA-3'
Poly(G) ₁₈	5'-CCCCCCCCCCCCCTATAGTGAGTCGTATTA-3'
Poly(A) ₁₈	5'-TTTTTTTTTTTTTTTTCCCTATAGTGAGTCGTATTA-3'
Poly(GU) ₉	5'-AACAAACAACAACAACCCCTATAGTGAGTCGTATTA-3'

Table 2 Oligonucleotides used for site-directed mutagenesis

Underlined nucleotide(s) indicate changes from the normal sequence.

Mutation	Primers for mutation
P41S*	5'-GAAACTAACAACTGACCCAACCTCAG-3' 5'-CTGAGGTTGGG <u>T</u> CAGTTGTAGTTTC-3'
T106P† (1)	5'-GAAGAGCCTTGGCC <u>C</u> CTGGTGCACCTG-3' 5'-CAGGTGCACCAAGG <u>G</u> CCCAAGGCTCTTC-3'
G107A† (4)	5'-GGCCCTGCTGCACCTGTC-3' 5'-GACAGGTGCAGCAGGGCC-3'
V110I, E112D† (3)	5'-CTGCACCTATCATAGATTCACCTTATG-3' 5'-CATAAGGTGAATCTATGATAGGTGCAG-3'
E117D, S118D† (5)	5'-CCTTATGGAGATCCATCAGCCCTG-3' 5'-CAGGGCTGATGGG <u>A</u> TCTGGATAAGG-3'
S120D† (6)	5'-GATCCCATCGACCCCTGAGGATG-3' 5'-CATCCTCAGGGT <u>C</u> GATGGGATC-3'
S129T, K130R† (2)	5'-GAATCCATTACCAGAGCAGTTGC-3' 5'-GCAACTGCTCTGGTAAATGGATTC-3'

* In this pair of oligonucleotides GST-CstF-64 RBD at amino acid 41 are mutated from a proline residue to a serine to create CstF-64 P41S.

† This set of oligonucleotide pairs allow for the different mutations on GST-CstF-64 RBD to be introduced sequentially to obtain the CstF-64 9AA construct. The number beneath the mutation corresponds to the order in which the mutations were performed.

by Midland Chemicals (Midland, TX, U.S.A.). ³²P-labelled poly(U)₁₈ substrate was synthesized by including 2 μl of [³²P]UTP (3000 Ci/mmol; PerkinElmer, Boston, MA, U.S.A.). RNAs were then phenol/chloroform ratio extracted followed by ethanol precipitation. Non-radiolabelled RNA ribopolymers were quantified spectrophotometrically by measuring absorbance at 260, 280 and 320 nm. Radiolabelled RNA substrate was washed in 10% (w/v) trichloroacetic acid after extraction and precipitation, then c.p.m. (counts per minute) were measured on a Wallac 1409 scintillation counter (PerkinElmer).

Site-directed mutagenesis

Site-directed mutagenesis was used to create two mutant constructs of the GST-CstF-64 RBD fusion protein using the QuikChange II Site-Directed Mutagenesis kit (Stratagene, La Jolla, CA, U.S.A.). The first construct mutated amino acid 41 of CstF-64 RBD from a proline to a serine residue ('CstF-64 P41S'; Table 2). A second construct included mutations for nine amino acids downstream of the RBD (T106P, G107A, V110I, E112D, E117D, S118D, S120D, S129T and K130R, 'CstF-64 9AA mutant'; Table 2).

UV-cross-linking and substrate competition assay

The UV-cross-linking assay used 0.2 μg (4.2 pmol) of either GST-CstF-64 RBD or GST-τCstF-64 RBD with one of two

amounts of ³²P-labelled poly(U)₁₈ substrate: 6.8 and 17 pmol. Simultaneously, competing amounts of one of five non-radioactive competitor RNAs [poly(U)₁₈, poly(G)₁₈, poly(A)₁₈, poly(C)₁₈ or poly(GU)₉] were added in the range 0.84–168.2 pmol in a total reaction volume of 20 μl in Buffer D (20 mM Hepes, pH 7.9, 1.5 mM MgCl₂, 10 mM KCl and 0.5 mM dithiothreitol) [43]. Binding of [³²P]poly(U)₁₈ was competed for by each non-radioactive ribopolymer; hence all affinities were measured relative to poly(U)₁₈. Reactions were incubated at 30 °C for 30 min and then exposed to 1.0 J/cm² of ultraviolet light in a CL-1000 UV cross-linker (UVP, Upland, CA, U.S.A.), followed by incubation with 10 units of RNase ONE™ (Promega, Madison, WI, U.S.A.) or RNase Cocktail™ (Ambion, Austin, TX, U.S.A.) at 37 °C for 15 min per reaction. Half of each reaction was boiled in SDS/PAGE loading buffer and separated by SDS/12.5%-PAGE. The gels were stained with Coomassie Brilliant Blue R-250, destained extensively in gel fix [20% (v/v) methanol and 10% (v/v) acetic acid], dried, and exposed to a storage phosphor screen (Molecular Dynamics, Sunnyvale, CA, U.S.A.) for band density quantification. Each condition was tested in triplicate and repeated at least three times.

Data analysis

Signal density data from the UV-cross-linking assays were analysed using Microsoft Excel spreadsheet program (Microsoft, Redmond, WA, U.S.A.). Binding inhibition constants (*K_i*) were extrapolated from Dixon plots for competitive inhibition as follows:

$$\frac{1}{v} = \left(\frac{K_m}{V_{max}[S]} \right) \left(1 + \frac{[I]}{K_i} \right) + \frac{1}{V_{max}}$$

where *v* is the reaction velocity, *S* is the substrate concentration, *I* is the inhibitor concentration, *V_{max}* is the rate of maximum turnover, *K_m* is the substrate concentration at half-maximal velocity, and *K_i* is the competitive inhibition constant. This method is useful since it provides a graphical means of determining *K_i* for competitive inhibitors [44,45]. At the point of intersection, 1/*v* and [I] are the same for both lines (see Figure 3 and [44]). Since the inhibition is competitive, *V_{max}* is also the same for both lines. Hence, the equation can be reduced to:

$$\frac{1}{[S_1]} \left(1 + \frac{[I]}{K_i} \right) = \frac{1}{[S_2]} \left(1 + \frac{[I]}{K_i} \right)$$

which is only true if either *S*₁ = *S*₂ or [I] = −*K_i*. Thus [I] at the point of intersection equals −*K_i*. In this case, *K_i* is equivalent to *K_d*, and we will refer to these values henceforth as *K_d* [44]. ANOVA statistical analyses and post tests were conducted using GraphPad InStat version 3.0 (GraphPad Software, San Diego, CA, U.S.A.).

RESULTS

Because we were concerned that binding equilibria might change dramatically in a two-phase 'partitioning' assay such as electrophoretic mobility-shift assay [46], and that a filter binding assay might not have enough sensitivity to discriminate binding differences, we chose to establish a UV-cross-linking assay for the analysis of RNA binding to CstF-64 RBDs. The UV-cross-linking assay has the advantages of high sensitivity, the complexes could be 'frozen' by the UV-cross-linking step, thereby eliminating concerns over stability of the complex, and the technique has been well established in the analysis of CstF-64 (cf. [23,24,47,48]). We also chose to use the N-terminal RNA-binding portions of

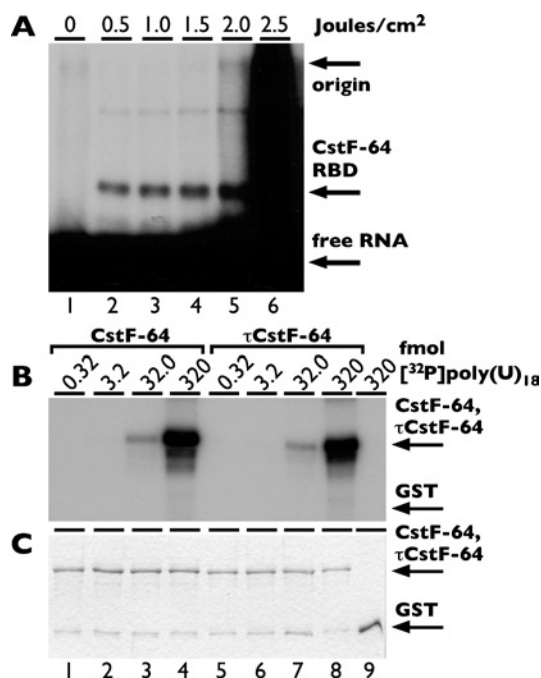


Figure 2 Conditions of UV cross-linking and RNA binding

(A) GST–CstF-64 RBD was bound to ^{32}P -labelled SV40 late pre-mRNA [24], exposed to increasing amounts of UV light from 0 J/cm^2 (lane 1) to 2.5 J/cm^2 (lanes 2–6), and separated by SDS/10% PAGE, dried, and analysed by autoradiography. Arrows indicate the mobility of GST–CstF-64 RBD (~45 kDa) and free RNA. (B) Increasing amounts of [^{32}P]poly(U) $_{18}$ from 0.32 to 320 fmol was incubated with 0.2 μg of either GST–CstF-64 RBD (lanes 1–4), GST– τ CstF-64 RBD (lanes 5–8), or GST (lane 9). After a 30 min incubation, the material was irradiated with 1.0 J/cm^2 of UV light and processed for analysis as indicated in the Experimental section. Arrows indicate the mobilities of GST–CstF-64 RBD, GST– τ CstF-64 RBD (~45 kDa) and GST (~26 kDa). (C) Coomassie Blue staining of SDS/PAGE from (B), indicating mobilities of GST–CstF-64 RBD, GST– τ CstF-64 RBD and GST. Note that the Coomassie-stained product at approx. 26 kDa in lanes 1–8 is a breakdown product of the GST-fusion proteins.

CstF-64 and τ CstF-64 in the binding assays because the full-length protein inhibits binding of the RBD to RNA [23,24], the CstF-64 RBD region has been well characterized in other RNA-binding experiments [27,49,50], and the full-length proteins are not expressed well in bacteria ([24,37] and results not shown).

Conditions for the UV-cross-linking/binding assay

To determine optimal conditions of cross-linking, we incubated 0.2 μg of GST-fused CstF-64 RBD with 10000 c.p.m. of ^{32}P -labelled SV40 (simian virus 40) late pre-mRNA substrate [24] at 30 °C for 30 min and then submitted the mixture to the UV fluxes indicated (Figure 2A). Mixtures were then separated by SDS/10% PAGE and visualized by autoradiography. In the absence of UV light, no ^{32}P -labelled GST–CstF-64 RBD was detected (Figure 2A, lane 1). Lanes 2–4 showed radioactive material at approx. 45 kDa, consistent with efficient cross-linking of ^{32}P -labelled RNA to the GST–CstF-64 RBD (lanes 2–4). However, there was little evidence of slower migrating radioactive material in those lanes (lanes 2–4), indicating minimal higher order cross-linking. At higher amounts of UV flux (lanes 5 and 6), there was considerable slower migrating radioactivity, probably resulting from protein–protein cross-linking and other higher order events. Based on these results, we performed all subsequent experiments at 1.0 J/cm^2 UV flux.

In order to establish the linear range of substrate concentrations, we tested binding of both GST–CstF-64 RBD and GST– τ CstF-64 RBD to a range of ^{32}P -labelled poly(U) $_{18}$ substrates

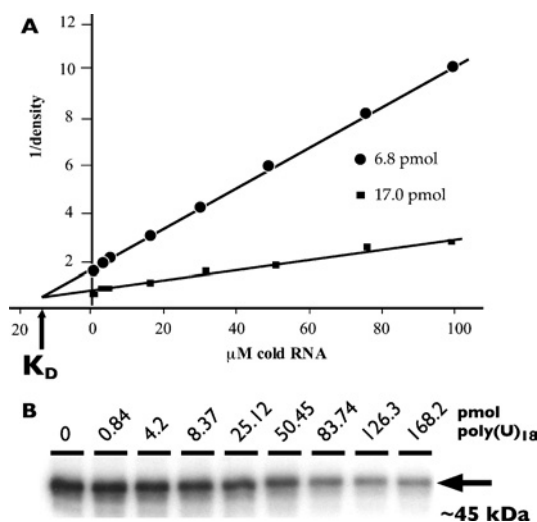


Figure 3 Competition/UV-cross-linking assay to determine relative affinities of RNA binding to the CstF-64 or τ CstF-64 RBDs

(A) Representative Dixon plot showing inhibition of UV cross-linking of the GST–CstF-64 RBD to 6.8 and 17 pmol of [^{32}P]poly(U) $_{18}$ by unlabelled poly(U) $_{18}$. See the Experimental section for details. (B) Representative autoradiograph of GST–CstF-64 RBD binding to [^{32}P]poly(U) $_{18}$ (17 pmol) in the presence of increasing amounts (in pmol as indicated) of unlabelled poly(U) $_{18}$. The arrow indicates the approximate mobility of the GST–CstF-64 RBD construct as determined by Coomassie Blue staining.

(Figure 2B). The amount of [^{32}P]poly(U) $_{18}$ used ranged from 0.32 to 320 fmol. Signal from UV-cross-linked RNA increased linearly for both GST–CstF-64 RBD (lanes 1–4) and GST– τ CstF-64 RBD (lanes 5–8). GST alone showed no RNA cross-linking even at 320 fmol of [^{32}P]poly(U) $_{18}$ (lane 9). Figure 2(C) shows Coomassie Blue staining of the gel to indicate amounts of protein loaded.

We also tested binding of [^{32}P]poly(U) $_{18}$ to GST–CstF-64 RBD with a range of 2–320 pmol of unlabelled poly(U) $_{18}$ in a competition assay [results not shown, but see Figure 3 and Supplementary Figure S.1 (<http://www.BiochemJ.org/bj/401/bj4010651add.htm>)]. Based on these results, we chose 6.8 and 17 pmol of [^{32}P]poly(U) $_{18}$ as the appropriate concentrations to use in competition assays for Dixon analysis of the K_i because they flanked the apparent K_d [44].

CstF-64 and τ CstF-64 have different affinities for RNA elements

Although CstF-64 and τ CstF-64 share a great deal of sequence identity, variations in and around the RBD may account for the differences in binding to RNA elements (Figure 1). To measure these differences, we designed a competition and UV-cross-linking assay to determine the relative affinities of GST-fusion proteins containing the CstF-64 and τ CstF-64 RBDs for simple ribopolymers [poly(U), poly(A), poly(G), poly(C) and poly(GU)]. These ribopolymers were chosen as simple models of more complex sequences that are often found downstream of authentic polyadenylation sites [33,51]. Recombinant CstF-64 and τ CstF-64 RBDs were incubated with either 6.8 or 17 pmol of ^{32}P -labelled poly(U) $_{18}$. To determine relative affinity for different ribopolymers, increasing amounts of non-radioactive poly(U) $_{18}$, poly(A) $_{18}$, poly(G) $_{18}$, or poly(C) $_{18}$ were included in the incubations as competitors. To separate bound from unbound RNA, the mixtures were allowed to equilibrate, and then cross-linked by exposure to UV light, treated with ribonuclease and separated by SDS/PAGE (Figure 3B). Radioactivity bound by either the CstF-64 or τ CstF-64 RBD was quantified, and the binding inhibition constant (K_i) was determined by Dixon plot analysis [44]. Figure 3(A) shows

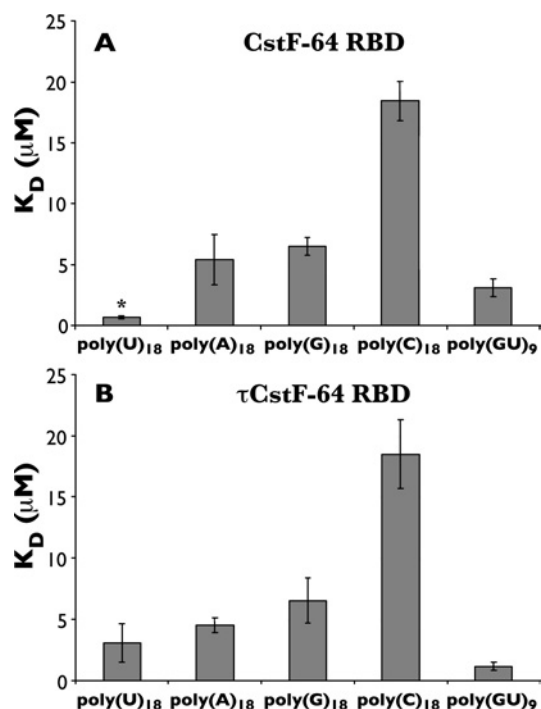


Figure 4 RNA-binding affinities of CstF-64 and τ CstF-64 for ribopolymers [poly(U)₁₈, poly(A)₁₈, poly(G)₁₈, poly(C)₁₈ and poly(GU)₉]

(A) K_d values (micromolar) relative to poly(U)₁₈ of the GST-CstF-64 RBD for poly(U)₁₈ ($K_d = 0.66 \pm 0.13 \mu\text{M}$), poly(A)₁₈ ($K_d = 5.4 \pm 2.05 \mu\text{M}$), poly(G)₁₈ ($K_d = 6.51 \pm 0.74 \mu\text{M}$), poly(C)₁₈ ($K_d = 18.44 \pm 1.62 \mu\text{M}$) or poly(GU)₉ ($K_d = 3.09 \pm 0.74 \mu\text{M}$). Poly(U)₁₈ ($P < 0.001$) and poly(C)₁₈ ($P < 0.001$) bindings were statistically significant according to ANOVA analysis with Tukey-Kramer multiple comparisons post-test. (B) τ CstF-64 has preferential binding to poly(GU)₉ compared with the other polymers [poly(U)₁₈, $K_d = 3.04 \pm 1.56 \mu\text{M}$; poly(A)₁₈, $K_d = 4.49 \pm 0.61 \mu\text{M}$; poly(G)₁₈, $K_d = 6.52 \pm 1.85 \mu\text{M}$; poly(C)₁₈, $K_d = 18.49 \pm 2.82 \mu\text{M}$; poly(GU)₉, $K_d = 1.14 \pm 0.34 \mu\text{M}$]. Binding to poly(C)₁₈ was statistically significantly different from all other values according to ANOVA analysis ($P < 0.001$).

an example plot comparing inhibition of [³²P]poly(U)₁₈ binding and cross-linking to the CstF-64 RBD by unlabelled poly(U)₁₈. Using this assay, we were able to measure K_i , which is equivalent to the dissociation constant (K_d) relative to poly(U) under the conditions described here [52]. We will henceforth refer to the values obtained as K_d values.

When CstF-64 was incubated with each of the different competitor ribopolymers ([poly(U)₁₈, poly(A)₁₈, poly(G)₁₈ or poly(C)₁₈] in the presence of substrate [³²P]poly(U)₁₈, CstF-64 had the highest affinity for poly(U)₁₈ ($K_d = 0.66 \pm 0.13 \mu\text{M}$) compared with the other ribopolymer competitors tested (Figure 4A). This is consistent with several previous studies that suggested that CstF-64 bound to U-rich elements [24,27,50,53,54]. When tested for the same ribopolymer competitors, τ CstF-64 did not bind as strongly to poly(U) ($K_d = 3.04 \pm 1.56 \mu\text{M}$) as did CstF-64 (Figure 4B). When poly(G)₁₈ and poly(A)₁₈ were tested for binding to τ CstF-64, they showed similar binding affinities as CstF-64 ($K_d \approx 4.5\text{--}6.5 \mu\text{M}$) (Figure 4B). Finally, both CstF-64 and τ CstF-64 bound poly(C)₁₈ with the weakest affinity ($K_d \approx 18.5 \mu\text{M}$) (Figures 4A and 4B).

Previously, the DSE has been described as either U-rich or GU-rich [24,27,50,54–57]. Our results in Figure 4(B) showed that CstF-64 RBD bound with highest affinity to poly(U)₁₈, as did τ CstF-64 RBD, although its binding was weaker than that of CstF-64. To test the affinity with which elements that contain guanosine might bind to either CstF-64 or τ CstF-64, we tested a polymer of

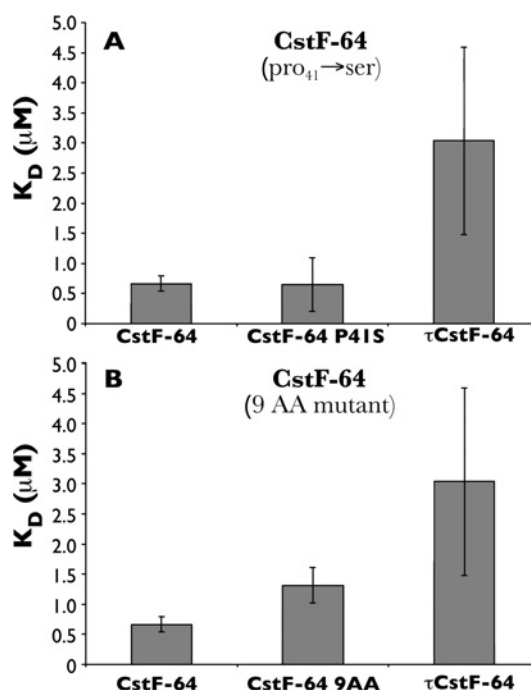


Figure 5 Contributions of different amino acids to CstF-64 and τ CstF-64 binding preferences to poly(U)

(A, B) K_d values of GST-CstF-64 RBD and mutants for poly(U)₁₈. (A) Amino acid 41 does not contribute to poly(U)₁₈ RNA-binding specificity. When the GST-CstF-64 RBD construct is mutated at amino acid 41 from a proline residue to a serine residue (CstF-64 P41S), there is no change in binding towards poly(U)₁₈ ($K_d = 0.64 \pm 0.45 \mu\text{M}$; no statistical difference is noted according to ANOVA analysis). (B) Amino acids downstream of the C-terminal α -helix contribute to RNA-binding specificity. When the GST-CstF-64 RBD construct was mutated at nine amino acids C-terminal to α_3 -helix (CstF-64 9AA), there was a decrease in binding to poly(U)₁₈ ($K_d = 1.31 \pm 0.29 \mu\text{M}$). This changes the binding of CstF-64 RBD to resemble that of τ CstF-64 for poly(U)₁₈.

(GU)₉. We determined a binding affinity of CstF-64 for poly(GU)₉ ($K_d = 3.09 \mu\text{M} \pm 0.74$; Figure 4A), slightly lower than previously reported by Deka et al. [49]. More interestingly, we found that τ CstF-64 bound poly(GU)₉ even more efficiently ($K_d = 1.14 \mu\text{M} \pm 0.34$) than did CstF-64 (Figure 4B), suggesting differences between the paralogues with regard to their affinity for both U-rich and GU-rich sequences.

Changes in amino acid 41 do not affect RNA binding

The CstF-64 and τ CstF-64 RBDs differ at amino acid 41 where CstF-64 has a proline residue and τ CstF-64 has a serine residue. To test whether this particular amino acid difference is an important determinant of RNA binding, site-directed mutagenesis was conducted on the CstF-64 construct to mutate the proline residue at position 41 to a serine residue. We saw no significant change in binding to poly(U)₁₈ in the presence of substrate [³²P]poly(U)₁₈ upon changing Pro-41 (Figure 5A; CstF-64, $K_d = 0.66 \pm 0.13 \mu\text{M}$) to a serine residue (Figure 5A; CstF-64 P41S, $K_d = 0.64 \pm 0.45 \mu\text{M}$). This indicates that the amino acid at position 41 does not contribute significantly to the differences in RNA binding seen between CstF-64 and τ CstF-64.

Elements downstream of the RBD may affect RNA-binding specificity

In addition to the proline \rightarrow serine mutation at position 41, CstF-64 and τ CstF-64 also differ at nine amino acids downstream of α_3

of the RBD (Figure 1). This newly described helix is thought to occlude the RBD, but may not participate directly in RNA selection [49]. To determine whether these amino acids affect affinity for poly(U), site-directed mutagenesis was used to create a GST–CstF-64 RBD construct with mutations at these nine amino acids [CstF-64(9AA); Table 2B] so that it resembled τ CstF-64 at these positions. The binding affinity of the CstF-64(9AA) mutant was tested as previously described (see Figures 3A and 3B) for poly(U)₁₈, as this substrate, showed the highest affinity and most significant difference in binding to normal CstF-64 and τ CstF-64. We found that the binding affinity of CstF-64(9AA) for poly(U)₁₈ was lower ($K_d = 1.31 \pm 0.29 \mu\text{M}$) than that of the wild-type CstF-64 (CstF-64, $K_d = 0.66 \pm 0.13 \mu\text{M}$), and began to approach that of τ CstF-64 (Figure 5B, $K_d = 3.04 \pm 1.56 \mu\text{M}$). This suggested that at least part of the discrimination of CstF-64 and τ CstF-64 for poly(U) is determined by the region downstream of the third α -helix (α_3).

DISCUSSION

We are interested in understanding the role of τ CstF-64 in polyadenylation during spermatogenesis. More specifically, we are studying this process in male germ cells because they offer a unique system wherein the mitotic cells express CstF-64, while the meiotic and post-meiotic cells express the paralogous τ CstF-64. Both paralogues are essential for spermatogenesis to proceed in an uninterrupted manner (B. Dass, R. A. Hess, K. Carnes, S. Tardif, H. M. Weitlauf and C. C. MacDonald, unpublished work). Recent reports indicate that the testis has both a higher incidence of alternative polyadenylation, suggesting differences in polyadenylation site choice, and a greater use of non-AAUAAA polyadenylation signals, suggesting differences in the polyadenylation machinery [33]. This led us to study whether τ CstF-64 contributes to either of these differences in polyadenylation. Therefore, as a first step, we proposed to examine the RNA-binding capabilities of τ CstF-64 to simple RNA substrates and compare them with the binding capabilities of CstF-64. To test the hypothesis that CstF-64 and τ CstF-64 bound to different RNA elements, we used a UV-cross-linking/competition assay for five ribopolymers and then used a Dixon plot to analyse the necessary parameters to determine K_d in this system [44,52]. Thus we obtained the K_d values of GST–CstF-64 RBD and GST– τ CstF-64 RBD for poly(U) and the K_d relative to poly(U)₁₈ for poly(A)₁₈, poly(G)₁₈, poly(C)₁₈ and poly(GU)₉.

Using this assay, we found that the CstF-64 RBD bound to synthetic poly(U)₁₈ RNA with much greater efficiency than to other synthetic ribohomopolymers (Figure 4A). The CstF-64 RBD bound with similar efficiency to poly(G) and poly(A), and bound quite weakly to poly(C). Binding of the CstF-64 RBD to a poly(GU) ribopolymer was intermediate between those of poly(U) and poly(G) [Figure 4A, poly(GU)₉]. From these binding affinities, we might predict that DSEs in natural polyadenylation sites would be U-rich. Use of U-rich DSEs has been supported in some studies [24,53,54], while in other studies GU-rich elements were implicated [56,58,59]. When *in vitro* RNA SELEX (systematic evolution of ligands by exponential enrichment) selection–amplification experiments were performed by two different groups, both found evidence supporting binding to GU-rich elements, although with a notable U bias [50,57]. More recently, Varani and co-workers [27,49] noted that binding of the CstF-64 RBD to a GU-rich synthetic RNA was enhanced by inclusion of a UU dinucleotide. Similarly, Chou et al. [53] suggested that a “four out of five base uridylylate tract” was sufficient to restore efficient polyadenylation to an SV40 late mRNA polyadenylation

signal. Therefore a UU dinucleotide is most likely an important part, although not necessarily the only part, of the DSE.

Because of the very weak binding to poly(C), another prediction might be that both the CstF-64- and τ CstF-64-binding sites would have a paucity of C residues, and that C residues could delimit binding sequences. These results are supported by solution data from Pérez Cañadillas and Varani [27] using the isolated CstF-64 RBD. In further support, SELEX experiments using GST–CstF-64 RBDs to select RNA aptamers resulted in RNAs that were U-rich and GU-rich, flanked by C residues [50]. In contrast, in similar SELEX experiments using purified CstF complex (containing all three subunits) from calf thymus or HeLa cells RNA aptamers rich in G, U and C were selected [57]. This suggests that the binding specificity of the intact CstF complex depends on more than the simple binding affinity of the CstF-64 RBD for RNA, and could be different when the intact CstF-64 is in complex with the other polyadenylation factors.

It seemed that the binding affinity of CstF-64 for any of the ribopolymers was relatively low (micromolar range). Nevertheless, results of Deka et al. [49] support this range ($K_d \approx 7\text{--}14 \mu\text{M}$). This suggests that the affinity of CstF-64 for RNA probably is enhanced during the substrate recognition step of polyadenylation by other interacting factors. This notion is supported by older observations that the interaction of CstF and CPSF was co-operative, and that CstF-64 binding required the presence of CPSF ([22] and C. C. MacDonald, unpublished work). It is further likely that the proximity of polyadenylation factors to the nascent pre-mRNA during transcription [60,61] contributes greatly to the specificity of substrate recognition by increasing the local presentation of RNA elements to the CstF and CPSF. We therefore have great enthusiasm for the model suggested by Pérez Cañadillas and Varani [27] that a feature of CstF-64, possibly the third α -helix (α_3), blocks casual RNA binding, but that the block is relieved when both CstF and CPSF interact with each other and with an authentic pre-mRNA substrate.

Early anecdotal references pointed out that the canonical AAUAAA polyadenylation signal was often not present in male germ cell mRNAs [62–67]. More than likely, this is evidence of a larger incidence of alternative polyadenylation during spermatogenesis [32,34–36]. A possible mechanism for this was revealed in a recent study of RNA sequence elements near sites of polyadenylation of mRNAs that were expressed during different stages of male germ cell development [33]. This study revealed significant differences in polyadenylation control elements between premeiotic and post-meiotic male germ cells, further supporting the view that some of the mechanisms of polyadenylation might be altered in those cells. We believe that differences in the RBDs of CstF-64 and τ CstF-64 might account for some of these differences [32,37]. As detailed, when compared with each other, the CstF-64 and τ CstF-64 RBDs bound with equal affinity to poly(G), poly(A) and poly(C) (Figure 4). There were differences, however, in affinities for poly(U): CstF-64 bound four to five times more avidly to poly(U) than did τ CstF-64 (Figure 4). In fact, the K_d of τ CstF-64 for poly(U) was not significantly different than its K_d for poly(G) or poly(A), suggesting that it had little selectivity for poly(U) over these two other ribohomopolymers. In contrast, the binding affinities of CstF-64 and τ CstF-64 to poly(GU) were reversed: τ CstF-64 bound poly(GU) nearly three times as well as did CstF-64, suggesting that poly(GU) might be a candidate for τ CstF-64-specific DSEs in germ cell mRNAs.

According to Allain and colleagues [28,68], even though they are distant from the RNA-binding surface, loops of RRM-type RBDs may contribute to additional RNA contact for extra specificity. Therefore we looked at amino acid sequences in or around the RBDs that might contribute to the differences in binding to

poly(U). When we compared amino acid sequence for CstF-64 and τ CstF-64 there was a single difference (proline \rightarrow serine) at amino acid 41 in the region between α_1 -helix and β_2 -sheet (Figure 1). This difference is conserved in orthologues of τ CstF-64 cloned from 15 mammalian species ([42,69] and B. Dass, unpublished work). However, site-directed mutagenesis on the CstF-64 RBD at this location (CstF-64 RBD-P41S, Figure 1) did not display any change in binding to poly(U)₁₈ (Figure 5A). This suggests that position 41 in the RBD does not contribute greatly to the specificity of binding of either CstF-64 or τ CstF-64 to poly(U).

Another series of differences between CstF-64 and τ CstF-64 lay C-terminal to α_3 described by Pérez Cañadillas and Varani [27]; these nine amino acids are similarly conserved in τ CstF-64 orthologues from 15 species (B. Dass, unpublished work). Upon testing the affinity of the CstF-64 RBD with the nine amino acid mutations for poly(U) (CstF-64 RBD-9AA; Figure 5B), we found that these changes lowered the affinity of the CstF-64 RBD for poly(U)₁₈, causing it to more closely resemble the τ CstF-64 RBD in affinity. This suggests to us that the region C-terminal to α_3 has a direct effect on the specificity of RNA binding. Pérez Cañadillas and Varani [27] proposed that α_3 was not directly involved in RNA binding, but instead served as a 'gate' to occlude RNA binding unless caused to move. Our own data support at least part of this model: while it likely serves as a gate, amino acids C-terminal to α_3 also affect RNA discrimination. We do not yet have a model for which protein-RNA contacts are necessary for this discrimination.

These data begin to address one of the mechanisms that might contribute to the high incidence of alternative polyadenylation described in male germ cell mRNAs. Since CstF-64 and τ CstF-64 are expressed in different temporal patterns during spermatogenesis [40] and have different RNA-binding specificities (the present study), they may recognize different DSEs of pre-mRNA polyadenylation signals of male germ cell pre-mRNAs. Our recent work suggests that there are a complex series of changes to 3'-end elements used as spermatogenesis proceeds [33]. Some of these changes include a reduction in the specificity of DSEs. Furthermore, 3'-untranslated regions in germ cell mRNAs were found to be notably shorter than in somatic cells. Since we saw a reduction in the discrimination of τ CstF-64 for certain sequences such as poly(U), this suggests a model in which τ CstF-64 might be more promiscuous in its choice of binding sites, thus binding to sites more 5' in the nascent transcript.

We acknowledge the NIH (National Institutes of Health; 2 R01 HD37109-04) and a minority student supplement to R. R. M. We also acknowledge Daniel Hardy, James Hutson and the members of the MacDonald laboratory for their discussions and contributions, S. Sridhara, Robert Shaw and J. Andrew Hockert for comments on this paper, Joel Graber for sharing results before publication, and an anonymous reviewer for suggesting a more consistent method of data analysis.

REFERENCES

- Zhao, J., Hyman, L. and Moore, C. (1999) Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol. Mol. Biol. Rev.* **63**, 405–445
- Edmonds, M. (2002) A history of poly A sequences: from formation to factors to function. *Prog. Nucleic Acid Res. Mol. Biol.* **71**, 285–389
- Logan, J., Falck-Pedersen, E., Darnell, Jr, J. E. and Shenk, T. (1987) A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse β maj-globin gene. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8306–8310
- Connelly, S. and Manley, J. L. (1988) A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes Dev.* **2**, 440–452
- Edwards-Gilbert, G., Prescott, J. and Falck-Pedersen, E. (1993) 3' RNA processing efficiency plays a primary role in generating termination-competent RNA polymerase II elongation complexes. *Mol. Cell. Biol.* **13**, 3472–3480
- Vagner, S., Vagner, C. and Mattaj, I. W. (2000) The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3'-end processing and splicing. *Genes Dev.* **14**, 403–413
- Huang, Y. and Carmichael, G. C. (1996) Role of polyadenylation in nucleocytoplasmic transport of mRNA. *Mol. Cell. Biol.* **16**, 1534–1542
- Jackson, R. J. and Standart, N. (1990) Do the poly(A) tail and 3' untranslated region control mRNA translation? *Cell* **62**, 15–24
- Sachs, A. B., Sarnow, P. and Hentze, M. W. (1997) Starting at the beginning, middle, and end: translation initiation in eukaryotes. *Cell* **89**, 831–838
- Wickens, M., Anderson, P. and Jackson, R. J. (1997) Life and death in the cytoplasm: messages from the 3' end. *Curr. Opin. Genet. Dev.* **7**, 220–232
- Decker, C. J. and Parker, R. (1994) Mechanisms of mRNA degradation in eukaryotes. *Trends Biochem. Sci.* **19**, 336–340
- Decker, C. J. and Parker, R. (1995) Diversity of cytoplasmic functions for the 3' untranslated region of eukaryotic transcripts. *Curr. Opin. Cell Biol.* **7**, 386–392
- Gunderson, S. I., Beyer, K., Martin, G., Keller, W., Boelens, W. C. and Mattaj, I. W. (1994) The human U1A snRNP protein regulates polyadenylation via a direct interaction with poly(A) polymerase. *Cell* **76**, 531–541
- Lutz, C. S. and Alwine, J. C. (1994) Direct interaction of the U1 snRNP-A protein with the upstream efficiency element of the SV40 late polyadenylation signal. *Genes Dev.* **8**, 576–586
- Lutz, C. S., Murthy, K. G., Schek, N., O'Connor, J. P., Manley, J. L. and Alwine, J. C. (1996) Interaction between the U1 snRNP-A protein and the 160-kD subunit of cleavage-polyadenylation specificity factor increases polyadenylation efficiency *in vitro*. *Genes Dev.* **10**, 325–337
- Bagga, P. S., Arhin, G. K. and Wilusz, J. (1998) DSEF-1 is a member of the hnRNP H family of RNA-binding proteins and stimulates pre-mRNA cleavage and polyadenylation *in vitro*. *Nucleic Acids Res.* **26**, 5343–5350
- Gunderson, S. I., Polycarpou-Schwarz, M. and Mattaj, I. (1998) U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol. Cell* **1**, 255–264
- Veraldi, K. L., Arhin, G. K., Martincic, K., Chung-Ganster, L. H., Wilusz, J. and Milcarek, C. (2001) hnRNP F influences binding of a 64-kilodalton subunit of cleavage stimulation factor to mRNA precursors in mouse B cells. *Mol. Cell. Biol.* **21**, 1228–1238
- Venkataraman, K., Brown, K. M. and Gilmartin, G. M. (2005) Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. *Genes Dev.* **19**, 1315–1327
- Jenny, A. and Keller, W. (1995) Cloning of cDNAs encoding the 160 kDa subunit of the bovine cleavage and polyadenylation specificity factor. *Nucleic Acids Res.* **23**, 2629–2635
- Murthy, K. G. K. and Manley, J. L. (1992) Characterization of the multisubunit cleavage-polyadenylation specificity factor from calf thymus. *J. Biol. Chem.* **267**, 14804–14811
- Wilusz, J., Shenk, T., Takagaki, Y. and Manley, J. L. (1990) A multicomponent complex is required for the AAUAAA-dependent cross-linking of a 64-kilodalton protein to polyadenylation substrates. *Mol. Cell. Biol.* **10**, 1244–1248
- Takagaki, Y., MacDonald, C. C., Shenk, T. and Manley, J. L. (1992) The human 64-kDa polyadenylation factor contains a ribonucleoprotein-type RNA binding domain and unusual auxiliary motifs. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 1403–1407
- MacDonald, C. C., Wilusz, J. and Shenk, T. (1994) The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. *Mol. Cell. Biol.* **14**, 6647–6654
- Nagai, K., Oubridge, C., Jessen, T. H., Li, J. and Evans, P. R. (1990) Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature* **348**, 515–520
- Varani, G. and Nagai, K. (1998) RNA recognition by RNP proteins during RNA processing. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 407–445
- Pérez Cañadillas, J. M. and Varani, G. (2003) Recognition of GU-rich polyadenylation regulatory elements by human CstF-64 protein. *EMBO J.* **22**, 2821–2830
- Maris, C., Dominguez, C. and Allain, F. H. (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J.* **272**, 2118–2131
- Takagaki, Y. and Manley, J. L. (2000) Complex protein interactions within the human polyadenylation machinery identify a novel component. *Mol. Cell. Biol.* **20**, 1515–1525
- Richardson, J. M., McMahon, K. W., MacDonald, C. C. and Makhatadze, G. I. (1999) MEARA sequence repeat of human CstF-64 polyadenylation factor is helical in solution: a spectroscopic and calorimetric study. *Biochemistry* **38**, 12869–12875
- Calvo, O. and Manley, J. L. (2001) Evolutionarily conserved interaction between CstF-64 and PC4 links transcription, polyadenylation, and termination. *Mol. Cell* **7**, 1013–1023

- 32 MacDonald, C. C. and Redondo, J.-L. (2002) Reexamining the polyadenylation signal: were we wrong about AAUAAA? *Mol. Cell. Endocrinol.* **190**, 1–8
- 33 Liu, D., Brockman, J. M., Dass, B., Hutchins, L. N., McCarrey, J. R., MacDonald, C. C., Singh, P. and Graber, J. H. (2007) Systematic variation in mRNA 3'-processing signals during mouse spermatogenesis. *Nucleic Acids Res.*, in the press
- 34 Edwalds-Gilbert, G., Veraldi, K. L. and Milcarek, C. (1997) Alternative poly(A) site selection in complex transcription units: means to an end? *Nucleic Acids Res.* **25**, 2547–2561
- 35 Zhang, H., Lee, J. Y. and Tian, B. (2005) Biased alternative polyadenylation in human tissues. *Genome Biol.* **6**, R100
- 36 Wallace, A. M., Dass, B., Ravnik, S. E., Tonk, V., Jenkins, N. A., Gilbert, D. J., Copeland, N. G. and MacDonald, C. C. (1999) Two distinct forms of the 64,000 *M_r* protein of the cleavage stimulation factor are expressed in mouse male germ cells. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6763–6768
- 37 Dass, B., McMahon, K. W., Jenkins, N. A., Gilbert, D. J., Copeland, N. G. and MacDonald, C. C. (2001) The gene for a variant form of the polyadenylation protein CstF-64 is on chromosome 19 and is expressed in pachytene spermatocytes in mice. *J. Biol. Chem.* **276**, 8044–8050
- 38 Handel, M. A., Hunt, P. A., Kot, M. C., Park, C. and Shannon, M. (1991) Role of sex chromosomes in the control of male germ-cell differentiation. *Ann. N.Y. Acad. Sci.* **637**, 64–73
- 39 Handel, M. A., Park, C. and Kot, M. (1994) Genetic control of sex-chromosome inactivation during male meiosis. *Cytogenet. Cell Genet.* **66**, 83–88
- 40 Wallace, A. M., Denison, T., Attaya, E. N. and MacDonald, C. C. (2004) Developmental differences in expression of two forms of the CstF-64 polyadenylation protein in rat and mouse. *Biol. Reprod.* **70**, 1080–1087
- 41 Wang, P. J., McCarrey, J. R., Yang, F. and Page, D. C. (2001) An abundance of X-linked genes expressed in spermatogonia. *Nat. Genet.* **27**, 422–426
- 42 Dass, B., McDaniel, L., Schultz, R. A., Attaya, E. and MacDonald, C. C. (2002) The gene CSTF2T encoding the human variant CstF-64 polyadenylation protein τ CstF-64 is intronless and may be associated with male sterility. *Genomics* **80**, 509–514
- 43 Dignam, J. D., Lebovitz, R. M. and Roeder, R. G. (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**, 1475–1489
- 44 Dixon, M. (1953) The determination of enzyme inhibitor constants. *Biochem. J.* **55**, 170–171
- 45 Segel, I. H. (1976) *Biochemical Calculations: How to Solve Mathematical Problems in General Biochemistry*, 2nd edn, John Wiley and Sons, New York
- 46 Weeks, K. M. and Crothers, D. M. (1992) RNA binding assays for Tat-derived peptides: implications for specificity. *Biochemistry* **31**, 10281–10287
- 47 Edwalds-Gilbert, G. and Milcarek, C. (1995) Regulation of poly(A) site use during mouse B-cell development involves a change in the binding of a general polyadenylation factor in a B-cell stage-specific manner. *Mol. Cell. Biol.* **15**, 6420–6429
- 48 Wilusz, J. and Shenk, T. (1988) A 64 kDa nuclear protein binds to RNA segments that include the AAUAAA polyadenylation motif. *Cell* **52**, 221–228
- 49 Deka, P., Rajan, P. K., Pérez-Cañadillas, J. M. and Varani, G. (2005) Protein and RNA dynamics play key roles in determining the specific recognition of GU-rich polyadenylation regulatory elements by human Cstf-64 protein. *J. Mol. Biol.* **347**, 719–733
- 50 Takagaki, Y. and Manley, J. L. (1997) RNA recognition by the human polyadenylation factor CstF. *Mol. Cell. Biol.* **17**, 3907–3914
- 51 Graber, J. H., Cantor, C. R., Mohr, S. C. and Smith, T. F. (1999) *In silico* detection of control signals: mRNA 3'-end-processing sequences in diverse species. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14055–14060
- 52 Linden, J. (1982) Calculating the dissociation constant of an unlabelled compound from the concentration required to displace radiolabel binding by 50%. *J. Cyclic Nucleotide Res.* **8**, 163–172
- 53 Chou, Z. F., Chen, F. and Wilusz, J. (1994) Sequence and position requirements for uridylyte-rich downstream elements of polyadenylation signals. *Nucleic Acids Res.* **22**, 2525–2531
- 54 Legendre, M. and Gautheret, D. (2003) Sequence determinants in human polyadenylation site selection. *BMC Genomics* **4**, 7
- 55 Wilusz, J. and Shenk, T. (1990) A uridylyte tract mediates efficient heterogeneous nuclear ribonucleoprotein C protein-RNA cross-linking and functionally substitutes for the downstream element of the polyadenylation signal. *Mol. Cell. Biol.* **10**, 6397–6407
- 56 Weiss, E. A., Gilmartin, G. M. and Nevins, J. R. (1991) Poly(A) site efficiency reflects the stability of complex formation involving the downstream element. *EMBO J.* **10**, 215–219
- 57 Beyer, K., Dandekar, T. and Keller, W. (1997) RNA ligands selected by cleavage stimulation factor contain distinct sequence motifs that function as downstream elements in 3'-end processing of pre-mRNA. *J. Biol. Chem.* **272**, 26769–26779
- 58 McLauchlan, J., Gaffney, D., Whitton, J. L. and Clements, J. B. (1985) The consensus sequence YGTGTTY located downstream from the AAUAAA signal is required for efficient formation of mRNA 3' termini. *Nucleic Acids Res.* **13**, 1347–1368
- 59 Nussinov, R. (1986) TGTG, G clustering and other signals near non-mammalian vertebrate mRNA 3' termini: some implications. *J. Biomol. Struct. Dynam.* **3**, 1145–1153
- 60 Bentley, D. (1999) Coupling RNA polymerase II transcription with pre-mRNA processing. *Curr. Opin. Cell Biol.* **11**, 347–351
- 61 Minvielle-Sebastia, L. and Keller, W. (1999) mRNA polyadenylation and its coupling to other RNA processing reactions and to transcription. *Curr. Opin. Cell Biol.* **11**, 352–357
- 62 Meijer, D., Hermans, A., von Lindern, M., van Agthoven, T., de Klein, A., Mackenbach, P., Grootegoed, A., Talarico, D., Valle, G. D. and Grosveld, G. (1987) Molecular characterization of the testis specific c-abl mRNA in mouse. *EMBO J.* **6**, 4041–4048
- 63 Oppi, C., Shore, S. K. and Reddy, E. P. (1987) Nucleotide sequence of testis-derived c-abl cDNAs: implications for testis-specific transcription and abl oncogene activation. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8200–8204
- 64 Øyen, O., Myklebust, F., Scott, J. D., Cadd, G. G., McKnight, G. S., Hansson, V. and Jahnsen, T. (1990) Subunits of cyclic adenosine 3',5'-monophosphate-dependent protein kinase show differential and distinct expression patterns during germ cell differentiation: alternative polyadenylation in germ cells gives rise to unique smaller-sized mRNA species. *Biol. Reprod.* **43**, 46–54
- 65 Wingett, D., Reeves, R. and Magnuson, N. S. (1992) Characterization of the testes-specific pim-1 transcript in rat. *Nucleic Acids Res.* **20**, 3183–3189
- 66 Foulkes, N. S., Schlotter, F., Pévet, P. and Sassone-Corsi, P. (1993) Pituitary hormone FSH directs the CREM functional switch during spermatogenesis. *Nature* **362**, 264–267
- 67 Ravnik, S. E. and Wolgemuth, D. J. (1996) The developmentally restricted pattern of expression in the male germ line of a murine cyclin A, cyclin A2, suggests roles in both mitotic and meiotic cell cycles. *Dev. Biol.* **173**, 69–78
- 68 Steff, R., Skrisovska, L. and Allain, F. H. (2005) RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. *EMBO Rep.* **6**, 33–38
- 69 Huber, Z., Monarez, R. R., Dass, B. and MacDonald, C. C. (2005) The mRNA encoding τ CstF-64 is expressed ubiquitously in mouse tissues. *Ann. N.Y. Acad. Sci.* **1061**, 163–172

Received 18 July 2006/14 September 2006; accepted 10 October 2006
 Published as BJ Immediate Publication 10 October 2006, doi:10.1042/BJ20061097