Mechanisms of translational repression of the Smcp mRNA in round spermatids

Danielle L Cullinane*, Tamjid A Chowdhury* and Kenneth C Kleene

Department of Biology, University of Massachusetts Boston, 100 Morrissey Boulevard, Boston, Massachusetts 02125-3393, USA

Correspondence should be addressed to K C Kleene; Email: Kenneth.kleene@umb.edu

*(D L Cullinane and T A Chowdhury contributed equally to this work)

Abstract

The protamine 1 (Prm1) and sperm mitochondria-associated, cysteine-rich protein (Smcp) mRNAs exemplify a widespread pattern of mRNA-specific regulation of mRNA translation in post-meiotic spermatogenic cells, spermatids. Both mRNAs are transcribed and initially stored in free-mRNPs in early spermatids, and translated on polysomes in late spermatids. In this study, we demonstrate that the 5’- and 3’-UTRs and the 3’ terminus of the Smcp 3’-UTR are required for normal repression of the Smcp mRNA in transgenic mice. RNA affinity chromatography and mass spectrometry sequencing identified Y-box protein 2 (YBX2/MSY2) as the major protein that interacts with the 3’ terminus of the Smcp 3’-UTR and a Y-box recognition sequence, GCCACC, in the translation control element that is necessary for Prm1 mRNA repression. Depletion of YBX2 in Ybx2-null mice prematurely activates Prm1 and Smcp mRNA translation in early spermatids. This and previous findings suggest that the Smcp pre-mRNA is spliced and associates with YBX2 in the chromatoid body, and that repressed free-mRNPs are stored in the general cytoplasm. As YBX2 is the predominant protein in testis free-mRNPs, it likely represses many mRNAs in early spermatids. The mechanisms by which YBX2 represses the Smcp and Prm1 mRNAs are relevant to reproductive medicine because mutations in the human YBX2 gene correlate with abnormal protamine expression and male infertility.

Reproduction (2015) 149 43–54

Introduction

The developmental regulation of translational of specific mRNAs is important in controlling protein expression during the development of post-meiotic, haploid spermatogenic cells, spermatids, because transcription in late spermatids ceases due to chromatin remodeling (reviewed in Meistrich et al. (2003) and Kleene (2013)). This phenomenon is illustrated by the protamine 1 (Prm1) and sperm mitochondria cysteine-rich protein (Smcp) mRNAs, which are transcribed in round spermatids stored as translationally inactive messenger ribonucleoprotein particles (free-mRNPs) for 3 and 7 days, respectively, before translation begins in transcriptionally committed elongating and elongated spermatids (Kleene 1989, 2013). The initial block to mRNA translation avoids deleterious effects of premature PRM1 expression on male fertility (Lee et al. 1995).

mRNA-specific translational regulation usually involves cis-elements which bind trans-factors, RNA-binding proteins (RBPs) or small noncoding RNAs, which activate or repress translation (reviewed in Jackson et al. (2010)). About 20 RBPs and several miRNAs have been reported to be implicated in translational control in spermatids by studies of knockout mice and overexpression of RBPs (Kleene 2013, Kotaja 2014). However, RBPs and miRNAs often interact with huge numbers of mRNA targets and are expressed for prolonged periods in spermatogenic cells. This complexity creates uncertainties whether the effects of individual factors on target mRNAs are direct or indirect and whether the factor produces major or minor changes in translational activity (Kleene 2013). The functions of factors in translational regulation can be accurately defined by analyzing mutations in cis-elements that abrogate binding of factors. Unfortunately, this approach is rarely used because mutations are prone to negative results and the only system in which mutations can be studied in spermatids, transgenic mice, is expensive and laborious.

One of the atypical features of post-transcriptional gene expression in spermatids is the chromatoid body, a prominent, ~1 μm diameter, perinuclear RNP granule containing thousands of mRNA species, and a plethora of RBPs and noncoding RNAs with diverse functions in RNA biology (Meikar et al. 2014). Gene knockouts that
deplete chromatoid body constituents produce male infertility, suggesting that the chromatoid body coordinates regulatory processes that are necessary for spermatid development (Meikar et al. 2011). However, it is unclear whether the chromatoid body functions as a remodeling center, which regulates the interactions of RBPs and miRNAs with mRNAs, or a realization center, which carries out the consequences of these interactions. An especially relevant question here is whether translationally repressed free-mRNPs are formed and stored in the chromatoid body or formed in the chromatoid body and stored in the general cytoplasm (reviewed in Kleene & Cullinane 2011).

At present, the Prm1 and Smcp mRNAs are the only mRNAs in which multiple mutations have been analyzed in transgenic mice. A remarkable series of deletion and point mutations identify a 17-nucleotide (nt) translational control element (TCE) in the Prm1 3′-UTR immediately upstream of the AAUAAA polyadenylation signal as necessary for repression in step 7–9 spermatids (Zhong et al. 2001). However, the mechanism by which the TCE represses translation is unknown because a factor that binds the TCE has not been identified.

Previous studies using the GFP reporter suggest that the week-long repression of the Smcp mRNA in step 3–10 spermatids is regulated by multiple mechanisms involving both UTRs (Hawthorne et al. 2006, Bagarova et al. 2010). The Smcp 5′-UTR alone delays GFP expression until step 5, the Smcp 3′-UTR alone delays GFP expression until step 9, and a mutation in the Smcp 3′-UTR partially releases translational repression (Bagarova et al. 2010). Clearly, cis-elements and factors that repress the Smcp mRNA until step 11 spermatids remain to be identified.

This study continues our objectives of identifying the cis-elements and trans-factors that repress the Smcp mRNA. We have analyzed two new transgenes. The first transgene contains the Smcp 5′-UTR and 3′-UTR to test the proposition that both UTRs are necessary to delay translational activation until step 11. The second transgene mutates a conserved segment in the Smcp 3′-UTR downstream of the first AAUAAA polyadenylation signal (Chowdhury & Kleene 2012). The unusual position of this mutation is based on the evidence that elements which repress Prm1 mRNA translation are located at the 3′ terminus of the 3′-UTR, referred to as the 3T3U below (Giorigni et al. 2001, Zhong et al. 2001, RE Braun 2013, personal communication). We also use RNA affinity chromatography and mass spectrophotometry sequencing to demonstrate that Y-box protein 2 (YBX2/MSY2) binds the Prm1 TCE and the Smcp 3T3U, and demonstrate that the ybx2-null mutation results in premature activation of Prm1 and Smcp mRNA translation. Finally, a highly sensitive fluorescent in situ hybridization protocol reveals that natural and transgenic mRNAs with widely different translational activities are concentrated in the chromatoid body.

Materials and methods

Animal research

Protocols for the maintenance and usage of mice in this study were reviewed and approved by the University of Massachusetts Boston IACUC, Assurance # A3383-01, and are in accordance with the 2011 NIH ‘Guide for the Care and Use of Experimental Animals’. CD-1 and C57BL/6 mice were maintained on a 12-h light:12 h darkness cycle, provided with food and water ad libitum, and killed by CO2 hypoxia.

Construction of the S5GCS3 and G5GCS3-mut2 transgenic mice

The S5GCS3 transgene was constructed from G5GCS3 and S5GCS3 transgenes described previously (Hawthorne et al. 2006). Briefly, the G5GCS3 and S5GCS3 transgenes were digested with BsgI and AflIII, and the small fragment containing the Smcp 3′-UTR (S5) was ligated into the large fragment containing the Smcp 5′-UTR (S5) and EGFP-coding region (G5). The G5GCS3-mut2 transgene was constructed from the G5GCS3 and G5GCS3-g3−c1 transgenes in several steps. A Swa I site was inserted upstream of the first Smcp poly(A) signal with an overlap extension PCR in the G5GCS3 transgene (Higuchi et al. 1988). Next, the Swa I–Afl II fragment from the G5GCS3 transgene was inserted into the Swa I–Afl II sites of G5GCS3. Finally, the Swa I site was reversed to that of the natural Smcp 3′-UTR with a second round of overlap extension PCR. The sequences of both transgenes are presented in Supplementary data 1, see section on supplementary data given at the end of this article.

Transgenic mice were generated by pronuclear injection (University of Massachusetts Medical Center Transgenic and Knockout Animal Core Facility Worcester, MA, USA). The techniques for the breeding and genotyping of transgenic mice have been described previously (Bagarova et al. 2010).

Analysis of GFP expression in the squashes of seminiferous tubules

The stage of GFP expression was analyzed in the squashes of living spermatogenic cells (Kotaja et al. 2004, Bagarova et al. 2010). Adult or immature mice were killed with CO2 hypoxia, the testes were dissected out and the tunica albuginea was removed. The seminiferous tubules from adult testes were teased apart in PBS and visualized using a dissecting microscope and transillumination to identify tubule segments from adult mice of potential interest. The stages of spermatids were identified in squashes of 0.5 mm tubule segments by phase-contrast microscopy using an Olympus BX51 microscope equipped with a Plan Fluorite 100× phase objective (NA 1.3), 100 W mercury burner, and SPOT XPLORER monochrome camera, SPOT image processing software (Diagnostic Instruments, Sterling Heights, MI, USA). EGFP fluorescence was excited at 470 nm and emitted light at 525 nm was photographed at a manual setting of 3 s and γ = 1. Image J 1.45S was downloaded from the NIH (Bethesda, MD, USA) and used to quantify the levels of GFP expression in various cell types.
Sucrose and Nycodenz gradient analysis of polysomal loading

The cytoplasmic extracts of 21 dpp prepubertal and adult testes were prepared by dissecting the testes of adult and 21-dpp mice as described earlier. The subsequent procedures were carried out at 4 °C. The testes were homogenized in 300 μl HNM buffer (20 mM HEPES, pH 7.4, 0.1 M NaCl, 3 mM MgCl2) containing 0.5% Triton X-100 and 1 unit/μl RNAsin Plus (Promega Biotech, Kleene et al. 2010). The extract was centrifuged at 13 000 g for 2 min, and 250 μl of the supernatant was layered on either a 3.8 ml linear 15–40% sucrose gradient in HNM buffer (w/v) or a 3.8 ml of 20–60% (w/v) Nycodenz gradient prepared by layering 760 μl of 60, 50, 40, 30, and 20% Nycodenz (Accurate Scientific Chemical Corporation, Westbury, NY, USA) in HNM in polyallomer centrifuge tubes for the Beckman SW60 rotor. The sucrose gradients were centrifuged for 80 min at 125 000 g and ~0.4 ml fractions were collected onto 0.3 g guanine thiocyanate (Sigma–Aldrich) and Nycodenz gradients were centrifuged for 24 h at 140 000 g and ~0.2 ml fractions were collected onto 0.15 g guanidine thiocyanate. RNA was extracted from the pellet and fractions and the levels of the Ldhc, Smcp, and transgenic mRNAs in each fraction were determined by RT-qPCR as described previously (Bagarova et al. 2010, Kleene et al. 2010).

RNA affinity chromatography

Briefly, 20 μg 5′-biotinylated RNA (Sigma–Aldrich, Co.) was heated to 70 °C for 5 min in 400 μl binding buffer (20 mM HEPES, 3 mM MgCl2, 40 mM KCl, and 1 mM dithiothreitol, pH 7.6) containing protease inhibitor cocktail (Roche 11836170001) and 5% glycerol, slowly cooled, and incubated with ~500 μg extract proteins and 5 μg Escherichia coli tRNA for 30 min at 25 °C. The reactions were centrifuged with 2 μl of heparin (200 mg/ml) for 10 min and incubated with pre-washed streptavidin–agarose (Pierce 20347, Rockford, IL, USA) on a rotating disc for 2 h at 4 °C. After 5, 15 min, 1 ml washes with a binding buffer with protease inhibitors at 4 °C, the proteins were released by boiling for 5 min in 50 μl SDS sample buffer, separated by SDS–PAGE, and visualized by silver staining (Pierce 24600). The bands were excised and identified with trypsin digestion and mass spectrometry sequencing at the Taplin Mass Spectrometry Facility (Harvard Medical School, Boston, MA, USA).

Analysis of Prm1 and Smcp mRNA translation in Ybx2-null testis

Male and female mice bearing heterozygous WT and knockout alleles at the Ybx2 gene (Yang et al. 2007) were mated and Ybx2+/+, Ybx2+/−, and Ybx2−/− progeny were identified in DNA extracted from tail biopsies (Bagarova et al. 2010). The WT and knockout Ybx2 alleles were distinguished by PCR amplification using a reverse primer for both alleles, GCAGAACAGTGGGTGTGTT, and forward primers specific for the WT allele, GGAGGGGAAGGGGCAAT, and the knock-out allele, GCAGAACAGTGGGTGTGTT. The samples were denatured at 94 °C for 2 min, amplified with 27 cycles (94 °C 50 s, 57 °C 50 s, 72 °C 30 s) followed by 72 °C 7 min. The sequences of the primers and PCR program were kindly provided by Drs RM Schultz and S Medvedev.

Immunocytochemistry and RNA-fluorescent in situ hybridization

Stage II–VI seminiferous tubule segments were identified by transillumination and dissected in DEPC-treated PBS mechanically dispersed in DEPC-treated 100 mM sucrose and fixed and spread as dried-down preparations on slides dipped in 0.05% Triton X-100 and freshly prepared 4% formaldehyde (EM Sciences, Hatfield, PA, USA) (Kotaja et al. 2004). For RNA-fluorescent in situ hybridization (RNA-FISH), 20-nt oligo probe sets for Smcp coding region, the Smcp intron, the Gip coding region, and the Ldhc mRNA and introns were selected using the Stellaris Probe Designer at the Biosearch Technologies (Petaluma, CA, USA) website. The probe sets, consisting of 24–48 oligos, were fluorescently labeled with Quasar 570 or Quasar 670. The sequences of the probe sets and additional details of their selection are described in Supplementary data 2, see section on supplementary data given at the end of this article. The cells were permeabilized with 70% ethanol at room temperature (RT) for 1 h, rinsed twice in wash buffer (2× SSC, 10% deionized formamide (Ambion AM9342, Austin, TX, USA) for 5 min at 37 °C, and incubated overnight at 37 °C with a 1:50 dilution of each probe-set in a hybridization buffer (10% dextran sulfate (Sigma–Aldrich D8906) in wash buffer). After hybridization, the slides were washed 3 × for 30 min at 37 °C in wash buffer, rinsed with DEPC-treated PBS, and mounted in Prolong Gold Antifade reagent with DAPI (Life Technologies P36931).

For immunocytochemistry, dried-down preparations were rehydrated in PBS, blocked with 10% normal goat serum (5-1000, Vector Laboratories, Youngstown, OH, USA) for 20 min at RT, washed with DEPC PBS, and incubated with rabbit polyclonal antibody to mouse vasa homologue, MVH (Abcam, Eugene, OR, USA, ab13840, 1:200, overnight at 4 °C), or affinity-purified rabbit polyclonal antibody to Y-box protein 2, YBX2 (Yu et al. 2003) (1:200, 1 h, RT), washed with PBS, reacted with goat anti-rabbit secondary antibody (Alexa Fluor 488, A11008) or 594 (A11037) (1:500, 30 min, RT), washed in PBS, mounted, and counterstained with DAPI as described earlier.

For combined RNA-FISH and immunocytochemical studies, RNA-FISH was carried out through the final washing in PBS, followed by blocking and reaction with primary and secondary antibodies as described earlier. The cells were photographed using an Olympus BX51 microscope 100× plan fluoro objective equipped with Olympus filters, U-N31000 (excitation at 360 nm and emission at 460 nm) and U-N31004 (excitation at 560 nm and emission at 630 nm), or scanned at 0.8 μm using a Zeiss LSM 510 confocal microscope equipped with a Zeiss Plan-apo 63× NA oil objective. The wavelengths (nm) used for the confocal excitation and emissions of the fluoros in this study are as follows: DAPI, 405 and 460; Quasar 570, 547, and 570; Quasar 670, 644, and 670; Alexa fluor 488, 498, and 525; Alexa fluor 594, 594, and 617.

www.reproduction-online.org
Statistical analyses
Student’s paired, two-sided t-test was used to calculate P values for pixel intensities measured with ImageJ (InStat3, GraphPad Software, La Jolla, CA, USA).

Results
Developmental expression of GFP fluorescence in 5’GCS3 and G5GCS3-mut transgenic spermatids
Two transgenes were studied to clarify the functions of the Smcp UTRs in the developmental regulation of translation in spermatids. Both transgenes contain 511 nt of Smcp 5’ flanking region, which directs the transcription of Smcp–Gfp transgenic mRNAs in round spermatids at the same start site as the natural Smcp mRNA (Hawthorne et al. 2006). The segments of the transgenic mRNAs that are derived from Smcp and pEGFP are designated S and G respectively, and superscripts indicate the 5’-UTR, coding region and 3’-UTR. The first transgene, S5GCS3, contains the complete Smcp 5’-UTR and 3’-UTRs. The second, G5GCS3-mut2, contains the pEGFP 5’-UTR and the Smcp 3’-UTR, in which the 16 nt downstream of the first Smcp AAUAAA poly(A) signal has been replaced with the 17 nt downstream of the poly(A) signal in the pEGFP 3’-UTR. The full-length pEGFP 3’-UTR does not repress translation in round spermatids (Hawthorne et al. 2006, Bagarova et al. 2010). The structures of the transgenes and the sequences of the 3T3Us of the G5GCS3-mut1, and G5GCS3-mut2 transgenes are depicted in Fig. 1, and their complete sequences are presented in Supplementary data 1.

The developmental expression of GFP fluorescence was analyzed with phase-contrast and fluorescence microscopy of squashes of short segments of living seminiferous tubules (Fig. 2). The morphology of squashed spermatogenic cells is clearly visualized using phase-contrast microscopy (Kotaja et al. 2004). The visualization of weak GFP fluorescence in G5GCS3-mut2 round spermatids was facilitated by analyzing sexually immature 25–28 day post-partum (dpp) testes which lack intensely fluorescent-elongated spermatids.

GFP fluorescence was obvious in the nuclei and cytoplasm of step 3 round G5GCS3-mut2 spermatids and the cytoplasm of step 12 S5GCS3 elongated spermatids (Fig. 2). The uniform GFP fluorescence throughout the nuclei and cytoplasm of step 3 spermatids and the cytoplasm of step 12 spermatids is distinct from the faint mottled, cytoplasmic background fluorescence in nontransgenic spermatids and spermatocytes (Supplementary Fig. 1, see section on supplementary data given at the end of this article). The exclusion of EGFP from the nuclei of step 12 spermatids and the acrosomes of round spermatids has been noted previously (Ventelähl et al. 2000, Bagarova et al. 2010).

To document the developmental regulation of translational activity of the transgenic mRNAs more precisely, the intensities of GFP fluorescence of cell-free areas, spermatocytes, and spermatids were quantified using ImageJ. To reduce the effects of uneven illumination, the average pixel intensities of spermatids in close proximity to spermatocytes and cell-free areas were compared with Student’s two-sided paired t-test (Supplementary Fig. 1 and Table 1, see section on supplementary data given at the end of this article). The pixel intensities of spermatocytes in various stages of meiosis were indistinguishable from background of cell-free areas (P>0.26). The pixel intensities of step 1 G5GCS3-mut2 and step 11 S5GCS3 spermatids were on an average ~10% greater than those of spermatocytes or cell-free areas, and t-tests inconsistently produced P values <0.05. The pixel intensities of step 3 G5GCS3-mut2 and step 12 spermatids were ~1.5- to fourfold greater than those of spermatocytes and cell-free areas, and t-tests consistently produced P values <0.0001. Evidently, the initial repression of G5GCS3-mut2 and S5GCS3 mRNAs is leaky followed by translational activation in step 3 and step 12 spermatids respectively. The marked increase in GFP fluorescence in step 3 G5GCS3-mut2 and step

Figure 1 Structure and 3T3U sequences of transgenes. (A) The structures of the S5GCS3 and G5GCS3-mut2 transgenes are diagrammed. Segments derived from the pEGFP mRNA are green and segments derived from the Smcp mRNA are red. (B) Sequence of the 3’ termini of natural and mutant Smcp and Prm1 3’-UTRs. Y-box recognition sequences are highlighted yellow, canonical AAUAAA poly(A) signals are highlighted grey, and the 5’-UTR, in which the 16 nt downstream of the first Smcp AAUAAA poly(A) signal has been replaced with the 17 nt downstream of the poly(A) signal in the pEGFP 3’-UTR. The full-length pEGFP 3’-UTR does not repress translation in round spermatids (Hawthorne et al. 2006, Bagarova et al. 2010). The structures of the transgenes and the sequences of the 3T3Us of the G5GCS3-mut1, and G5GCS3-mut2 transgenes are depicted in Fig. 1, and their complete sequences are presented in Supplementary data 1.
that recovers equivalent proportions of RNA from each fraction (Kleene et al. 2010), and the levels of $S^G^C^S^3$ and $G^S^G^C^S^3$-mut2 mRNAs in each fraction were determined by RT-qPCR. The distribution of the transgenic mRNAs in each gradient was compared with those of the Smcp and Ldhc mRNAs. The $Ldhc$ mRNA, which encodes the testis-specific isoform of lactate dehydrogenase, provides a control for mRNA recovery and polysome integrity because it exhibits constant polysome loading in 21 dpp and adult testes (Bagarova et al. 2010, Kleene et al. 2010).

The $S^G^C^S^3$ and Smcp mRNAs displayed negligible levels of polysomal mRNA in sucrose gradients in 21 dpp testes, in which the most advanced cells are step 4

Sucrose gradient analysis of translational activity

To determine whether the differences in the developmental expression of GFP-fluorescence described above represent differences in translational activity, the translational activity of the $S^G^C^S^3$ and $G^S^G^C^S^3$-mut2 mRNAs was analyzed by sedimentation of cytoplasmic extracts from 21 dpp mice on sucrose gradients (Fig. 3). The cytoplasmic extracts were sedimented on sucrose gradients, fractions were collected from the bottom, RNAs were extracted using techniques that recover virtually identical proportions of RNA from each fraction, and the amounts of mRNA in each fraction were determined with RT-qPCR. The results are presented as graphs of the percentage of total mRNA on the gradient in each fraction. The symbols that identify the $Ldhc$, Smcp, $S^G^C^S^3$, and $G^S^G^C^S^3$-mut2 mRNAs are represented in each panel. The gradients demonstrate that the $S^G^C^S^3$ mRNA (A) is strongly repressed and that the $G^S^G^C^S^3$-mut2 mRNA (B) is active in step 1–4 spermatids. $S^G^C^S^3$ mRNA (Part A) is strongly repressed and the $G^S^G^C^S^3$-mut2 mRNA (Part B) is active.

Figure 2 Stage of first detection of GFP fluorescence in $S^G^C^S^3$ and $G^S^G^C^S^3$-mut2 transgenic round spermatids. Squashes of microdissected segments of seminiferous tubules were visualized by phase-contrast microscopy (left) to identify cell types and fluorescence microscopy (right) to monitor GFP expression. The contrast and brightness were adjusted to facilitate visualizing the morphology of spermatids and GFP fluorescence. The steps of spermatids were identified by cell associations and morphological criteria described by Russell et al. (1990). The arrow in the step 3 spermatid fluorescent cytoplasm points to an acrosome which appears dark because EGFP is excluded. The images demonstrate that dramatic increases in GFP expression occur in step 12 $S^G^C^S^3$ and step 3 $G^S^G^C^S^3$-mut2 spermatids.

12 $S^G^C^S^3$ spermatids was documented in three and five independent lines and founders respectively.

Figure 3 Sucrose gradients analysis of the translational activity of the $S^G^C^S^3$ and $G^S^G^C^S^3$-mut2 mRNAs in 21 dpp sexually immature testes. The cytoplasmic extracts were sedimented on sucrose gradients, fractions were collected from the bottom, RNAs were extracted using techniques that recover virtually identical proportions of RNA from each fraction, and the amounts of mRNA in each fraction were determined with RT-qPCR. The results are presented as graphs of the percentage of total mRNA on the gradient in each fraction. The symbols that identify the $Ldhc$, Smcp, $S^G^C^S^3$, and $G^S^G^C^S^3$-mut2 mRNAs are represented in each panel. The gradients demonstrate that the $S^G^C^S^3$ mRNA (A) is strongly repressed and that the $G^S^G^C^S^3$-mut2 mRNA (B) is active in step 1–4 spermatids. $S^G^C^S^3$ mRNA (Part A) is strongly repressed and the $G^S^G^C^S^3$-mut2 mRNA (Part B) is active.
spermatids (Kleene et al. 2010). In contrast, the G₅G₃S₃-mut2 mRNA exhibits high polysomal loading in 21 dpp testes, consistent with active translation and GFP expression. These findings demonstrate that the differences in the efficiency of translational initiation control the stages of first detection of GFP-fluorescence of the S₅G₃S₃ and G₅G₃S₃-mut2 mRNAs. These differences are supported by Nycodenz-gradient analysis shown in Supplementary Fig. 2, see section on supplementary data given at the end of this article, a technique that separates free-mRNPs and polysomes by differences in buoyant density (Kleene et al. 2010). In addition, sucrose and Nycodenz gradient analysis demonstrate that all four mRNA species exhibit substantial levels of polysomal mRNA in adult testis (Supplementary Fig. 2), implying that the S₅G₃S₃ and G₅G₃S₃-mut2 mRNAs are translationally active in elongated spermatids.

**RNA affinity chromatography of proteins binding to the 3′ termini of the Prm1, Smcp, and G₅G₃S₃-mut2 3′-UTRs**

To identify proteins that bind the 3′ termini of the Smcp, G₅G₃S₃-mut2, and Prm1 3′-UTRs, 5′ biotinylated RNA probes were incubated with testis protein extracts, treated with heparin to reduce nonspecific binding by electrostatic interactions, and protein–RNA complexes were captured with streptavidin–agarose resin (Fig. 4). After extensive washing, the bound proteins were eluted in SDS sample buffer, resolved by SDS–PAGE, and unique bands were identified by mass-spec sequencing.

Lanes 2, 4, and 6 display a single, prominent band at ~52 kDa. Mass spec sequencing of these bands identifies two Y-box proteins, Y-box protein 2, YBX2/MSY2 (NP_058571.2), and the long isoform of Y-box protein 3, YBX3L/MSY4 (NP_620817.2). YBX2 and YBX3L have nearly identical calculated molecular weights, 38.0 and 38.8 kDa, respectively, but migrate at ~52 kDa due to the anomalous mobility of Y-box proteins (Davies et al. 2000, Mastrangelo & Kleene 2000, Skabkin et al. 2006). Spectral counts reveal that YBX2 is more abundant than YBX3L. In addition, the average precursor intensities of the four most abundant peptides demonstrate that YBX2 is nine- to 14-fold more abundant than YBX3L.

The levels of the ~52 kDa band are drastically reduced in the incubations of testis extract with streptavidin beads without biotinylated RNA (lane 1). Mass spec sequencing reveals that this protein is β-tubulin 4A (NP_033477.2).

The segment of the Smcp 3′-UTR in lane 4 contains a Y-box 2 recognition sequence (YRS), AACAUUCU, that has been analyzed with more than 40 single and multiple point mutations (Bouvet et al. 1995, Giorgini et al. 2001, Bagarova et al. 2010, Chowdhury & Kleene 2012). These studies define YBX2 binding site as a 7 nt element, [ACGU][AC][CA][UC][AC]. Most permutations of alternative bases in brackets have little effect on YBX2 binding. The binding of YBX2 by the Prm1 TCE, lane 2, is eliminated by a GCCACCUC → GCCAGAU mutation, lane 3. This mutation has previously been demonstrated to abrogate YBX2 binding in vitro and translational repression by the Prm1 YRS in vivo (Giorgini et al. 2001). The binding of YBX2 to the Smcp 3T3U in lane 6 is unexpected because it does not contain an YRS, suggesting that YBX2 binds sequences that are not recognizable in the 7 nt degenerate YRS above.

The probe for the 3′ end of the G₅G₃S₃-mut2 3T3U binds two bands at ~51 and ~52 kDa (lane 5). The most abundant proteins in the 52 kDa band are YBX2 and YBX3L, but YBX2 is only 2.8-fold more abundant than YBX3L. The most abundant protein in the ~51 kDa band is the mouse homolog of the Lupus antigen (NP_001103615.1), which binds oligo(U) sequences (Allano et al. 2004), presumably the U-heptamer in the G₅G₃S₃-mut2 probe. Surprisingly, the bound proteins did not include ELAV1, another RBP with high affinity for U-rich elements and important functions in spermatogenesis (Chi et al. 2011).

The idea that YBX2 is the elusive factor that binds the Prm1 TCE YRS, GCCACCUC, is supported by u.v.-crosslinking assays shown in Supplementary Fig. 3, see section on supplementary data given at the end of this article. These assays demonstrate that YBX2 binds GCCACCUC in assays in which the TCE probe is crosslinked with protein before treatment with RNase T1. The observation that digestion with RNase T1 before crosslinking prevents detection of the complex...
with GCCACCU explains difficulties in the identification of YBX2 as the protein that binds the Prm1 TCE.

**Depletion of YBX2 results in premature recruitment of the Prm1 and Smcp mRNAs onto polysomes in round spermatids**

The evidence that YBX2 is the predominant protein that binds the Smcp 3T3U and Prm1 TCE suggests that depletion of YBX2 with the Ybx2-gene knockout will prematurely activate Prm1 and Smcp mRNA translation in round spermatids. Heterozygous Ybx2-knockout mice were mated (Yang et al. 2007), Ybx2<sup>−/−</sup> progeny were identified with PCR, and confirmed by the absence of YBX2 in immunocytochemistry in dried-down preparations of seminiferous tubules from 25 dpp Ybx2<sup>+/+</sup> and Ybx2<sup>−/−</sup> mice, Fig. 5A. DAPI-stained round spermatids are distinguished by round nuclei with bright central chromocenters. The YBX2-antibody stains the cytoplasm of Ybx2<sup>+/+</sup>-round spermatids with a higher concentration in a small perinuclear spot, the chromatoid body, in agreement with a previous immunohistochemical study of YBX2 in mouse testis (Oko et al. 1996). The YBX2-antibody did not stain Ybx2<sup>−/−</sup>-round spermatids.

The translational activity of the Prm1 and Smcp mRNAs was analyzed in sucrose gradients using the testis extracts of 25 dpp Ybx2<sup>+/+</sup> and Ybx2<sup>−/−</sup> mice, an age when the most advanced cells in the testis are step 9 spermatids and both mRNAs are strongly repressed (Braun et al. 1989, Fajardo et al. 1997, Kleene et al. 2010). Figure 5B reveals the absence of peaks of Prm1 and Smcp mRNAs in small polysomes in Ybx2<sup>+/+</sup> 25 dpp testes, consistent with strong repression, and the presence of obvious peaks in small polysomes in Ybx2<sup>−/−</sup> testes (Fig. 5C), consistent with active translation. The high levels of polysomal Ldhc mRNA in both gradients validate the integrity of the polysomes and suggest that YBX2 is not an important repressor of the Ldhc mRNA.

**Fluorescent in situ hybridization visualization of natural and transgenic mRNAs in chromatoid bodies in round spermatids**

The dramatic differences in translational activity of the 5′G<sup>5</sup>G<sup>5</sup>G<sup>5</sup>G<sup>5</sup>-mut2, Smcp, and Ldhc mRNAs in round spermatids provide an opportunity to analyze the localization of mRNAs in the chromatoid body, an RNP-granule that has been postulated to function in storing translationally repressed mRNAs in round spermatids (reviewed in Kleene & Cullinan 2011). We used an in situ hybridization protocol in which 24–48 tiled 20-base oligonucleotides specific for individual RNA species are 5′ end labeled with fluorochromes (Raj et al. 2008). This protocol achieves high sensitivity by combining excellent probe penetration, extensive target coverage, and a high ratio of fluorochromes to bases.

Confocal scanning microscopy with probes for the Smcp coding region and Ldhc mRNA (Fig. 6A and B) detected intense in situ hybridization signals in a ~ 1 μm diameter irregular perinuclear spot in dried down preparations from stage II–VI seminiferous tubules. The Smcp mRNA RNA-FISH co-localizes with immunocytochemical staining of mouse vasa homolog, MVH/DDX4 (Fig. 6H), an established marker for the chromatoid body (Yokota 2012, Meikar et al. 2014). The specificity of the 27 oligo
The Smcp mRNA probe is supported by findings that the subsets consisting the 14 odd and 13 even numbered oligos exhibit the identical patterns of hybridization (A Ortajo 2012, personal communication). In addition, RNAseq shows that the Ldhc and Smcp mRNAs are abundant in purified chromatoid bodies (Meikar et al. 2014).

The intense RNA-FISH staining of the translationally active Ldhc mRNA and the repressed Smcp mRNA in the chromatoid body implies that localization in the chromatoid body is independent of translational activity. The repressed $S^5C^3S^3$ and active $G^5C^5S^3$-mut2 mRNAs are also highly concentrated in the chromatoid body (Fig. 6C and D). The virtual absence of Gfp RNA-FISH signals in chromatoid bodies of nontransgenic mice demonstrate that the Gip probe is specific for Smcp–Gfp transgenic mRNAs (Fig. 6E). The derepressed Smcp mRNA in Ybx2-null testis is also concentrated in the chromatoid body (Fig. 6I).

The probes for the Smcp and Ldhc introns were used with the goal of detecting pulsatile pre-mRNA transcription in round spermatids in the absence of the intense chromatoid body RNA-FISH signal (Raj et al. 2006). Unexpectedly, the Smcp intron probe strongly stained the chromatoid body (Fig. 6F), while the Ldhc intron probe did not (Fig. 6G). The implications of these observations are considered in the Discussion.

Careful inspection of confocal microscope images does not reveal the presence of cytoplasmic mRNA (Fig. 6A, B, C, D, H and I). The relatively strong RNA-FISH signals in the chromatoid body can be explained by the fact that the chromatoid body represents <0.4% of the volume of the cytoplasm (Kleene & Cullinane 2011). This number can be used to calculate that if 5% of an mRNA was in chromatoid body and 95% of that mRNA was in the cytoplasm, the concentration of the mRNA would be 13-fold higher in the chromatoid body. The difference in mRNA concentration in the chromatoid body and cytoplasm is further exaggerated by 0.8 µm confocal optical sections which include many mRNAs in a small object, the chromatoid body, and fewer mRNAs in a thick object, the cytoplasm. Figure 6J demonstrates that the Smcp mRNA is detectable in the thin layer of cytoplasm surrounding round spermatid nuclei by conventional fluorescence microscopy, and that the Smcp mRNA co-localizes with YBX2, a marker for cytoplasmic free-mRNPs (Oko et al. 1996, Yang et al. 2007). The Smcp mRNA signal in nuclei may represent high levels of Smcp mRNA or the extremely diverse transcriptome in round spermatids (Soumillon et al. 2013).

Discussion

The evolutionarily unrelated Prm1 and Smcp mRNAs are the only mRNAs in which translational repression in round spermatids has been analyzed with multiple mutations in transgenic mice. These studies reveal important differences and similarities in mechanisms.
The principle difference is that the 3-day repression of the Prm1 mRNA is mediated entirely by the 3'-UTR TCE (Zhong et al. 2001), while the 8-day repression of the Smcp mRNA requires both the 5'-UTR and 3'-UTR. The 2-day repression by the Smcp 5'-UTR alone requires upstream reading frames which produce small polypeptides, whereas the 5-day repression by the Smcp 3'-UTR alone blocks initiation producing free-mRNPs (Bagarova et al. 2010). Additional mutations will be required to distinguish whether the 8-day repression by the Smcp 5'-UTR and 3'-UTR together is mediated by the upstream reading frames in the 5'-UTR and the block to initiation by the 3'-UTR, or interactions between the 3'-UTR and unidentified 5'-UTR elements that prolong the block to initiation.

The similarities are that the Prm1 and Smcp 3'-UTRs direct repression in free-mRNPs (Braun et al. 1989, Zhong et al. 2001, Bagarova et al. 2010). Significantly, Ybx2 binds the 3T3Us of the Prm1 and Smcp mRNAs and both mRNAs are prematurely activated in round spermatids by the Ybx2-null mutation. The importance of YRS position in the 3T3U is demonstrated by transgenes in which YRSs 3 and 16 nt upstream of the Prm1 AAUAAA poly(A) signal respectively repress translation strongly and partially (Fajardo et al. 1997, Giorgini et al. 2001, Zhong et al. 2001). In contrast, YRSs in the Prm1 and Smcp 3'-UTRs or > 34 nt upstream of the poly(A) signal in the 3'-UTR produce little or no repression (Zhong et al. 2001, Bagarova et al. 2010, RE Braun 2013, personal communication). The effects of YRS position imply that co-factors that bind the poly(A) signal, the poly(A) tail, or unidentified 3T3U elements strengthen repression by YBX2.

The importance of the 3T3U in mRNA repression in round spermatids is further supported by findings that repression is released by replacing the 16 nt downstream of first Smcp poly(A) signal with the 17 nt downstream of the pEGFP poly(A) signal. However, the pathway by which this mutation activates translation is unclear. The 23 nt segment extending from the AAUAAA poly(A) signal to the poly(A) site is known as the early SV40 polyadenylation signal (Kessler et al. 1986). As far as we are aware, the only proteins that bind this segment are CPSF30 and Wrd33, the canonical nuclear polyadenylation factors that bind AAUAAA (Chan et al. 2014). Therefore, the finding that the LA Lupus autoantigen (La) protein binds the G\(5\text{G}\text{C}\text{G}\text{G}-\text{mut2} 3\text{T3U} \text{was unexpected. The LA protein is necessary for early embryonic development (Park et al. 2006), binds oligo(U) sequences, and activates mRNA translation (Alfano et al. 2004), and its mRNA is expressed at high levels in testis (Carter & Sarnow 2000). The LA protein creates uncertainties whether the G\(5\text{G}\text{C}\text{G}\text{G}-\text{mut2} 3\text{T3U} \text{blocks the assembly of a complex that represses translation or stimulates translation by another pathway. All of the mRNA species studied here are concentrated in the chromatoid body in round spermatids, regardless of whether they are strongly repressed or translationally active. The prematurely translated Smcp mRNA in Ybx2-null round spermatids is also concentrated in the chromatoid body. Evidently, sequestration in the chromatoid body is not sufficient for prolonged repression, and sucrose gradient analysis suggests that free-mRNPs and polysomes are present in the general cytoplasm (reviewed in Kleene & Cullinane (2011)). The transitory localization of repressed mRNAs in the chromatoid body is clearly demonstrated by the first detection of the Tnp2 and Prm2 mRNAs simultaneously in the chromatoid body and the cytoplasm (Saunders et al. 1992, Fukuda et al. 2013).

The striking concentration of the Smcp intron in the chromatoid body was unexpected. However, noncanonical splicing in the cytoplasm is well-known (reviewed in Buckley et al. (2013)), and introns and splicing-factors are enriched in the chromatoid body (Biggiogera et al. 1990, Moussa et al. 1994, Meikar et al. 2014). The presence of the Ldhc mRNA and the absence of its introns in the chromatoid body are potentially related with the cessation of transcription of the Ldhc mRNA after meiosis (Tang & Goldberg 2012). One explanation is that the Ldhc mRNA is continuously present in germ cell RNA granules in spermatocytes and round spermatids. Another explanation is that the Ldhc mRNA is imported from the cytoplasm into the chromatoid body in round spermatids.

Two Y-box protein isoforms, YBX2 and YBX3L (MSY4), have been proposed to repress mRNA translation in round spermatids. The levels of both isoforms are maximal in round spermatids, followed by drastic decreases in elongating and elongated spermatids. These changes in levels correlate with the repression of many mRNAs in round spermatids and their subsequent activation in elongating and elongated spermatids (Oko et al. 1996, Davies et al. 2000). Our findings differ from previous ideas about the functions of YBX2 and YBX3L in mRNA repression in round spermatids.

YBX3L has been postulated to repress many mRNAs in round spermatids (Giorgini et al. 2002). However, YBX2 appears to be the critical isoform because the Ybx2-null mutation totally blocks spermatid development, whereas the Ybx3-null mutation has slight effects on spermatid development and male fertility (Lu et al. 2006, Yang et al. 2007). The more deleterious phenotype of the Ybx2-null mutation is likely related with observations that YBX2 is the predominant RBP in testis free-mRNPs (Herbert & Hecht 1999), and is much more abundant than YBX3L. YBX2 is also expected to bind more mRNAs because it binds single YRSs, whereas YBX3L forms a dimer with YBX2 that binds double YRSs (Davies et al. 2000). Single YRSs are much more numerous than double YRSs in repressed mRNAs in round spermatids (Chowdhury & Kleene 2012).

Another model proposes that YBX2 is a nonspecific RBP that selects pre-mRNAs for repression in the
cytoplasm by binding Y-box promoter elements in the nucleus (Yang et al. 2005a,b). This creates a propensity for YBX2 to bind nascent mRNAs, and YBX2 remains bound to the mRNA and represses translation after the mRNA is exported to the cytoplasm. This model is incompatible with evidence that Y-box proteins bind specific RNA sequences (Bouvet et al. 1995, Giorgini et al. 2001, Chowdhury & Kleene 2012, Wei et al. 2012) and exhibit no specificity for Y-box promoter elements (Zasedateleva et al. 2002, Dolfini & Mantovani 2013). Significantly, mutations in the Prm1 and Smcp mRNAs consistently demonstrate that mRNA sequences, not promoters, control repression in round spermatids (Braun et al. 1989, Schmidt et al. 1999, Zhong et al. 2001, Bagarova et al. 2010).

Understanding of the translational regulation of the Smcp mRNA is shown in Fig. 7. A bare-bones model maintains that the Smcp pre-mRNA is transcribed in the nucleus and exported through nuclear pores to the chromatoid body, where the pre-mRNA is spliced to produce the mature mRNA. YBX2 is predicted to bind 3T3U YRSs in the chromatoid body, because YBX2 is concentrated in the chromatoid body and undetectable in nuclei (Oko et al. 1996). The resulting mRNPs are exported to the cytoplasm and stored until the complex disassembles activating translation in step 12. The actual mechanism likely involves additional nuclear and cytoplasmic factors that facilitate the initial binding of YBX2 and its later dissociation from the YRS (Herbert & Hecht 1999, Zhong et al. 1999, Tsai-Morris et al. 2004, Matsumoto et al. 2005). Repression also likely requires factors that bind unidentified 3T3U elements and prevent binding of free-mRNPs to the small ribosomal subunit. As transcription, splicing, polyadenylation, and capping are normally tightly coupled in the nucleus (Bentley 2014), splicing in the chromatoid body creates uncertainty whether the 7-methylguanosine 5′ cap and 3′ poly(A) tail are added to pre-mRNAs in the nucleus or the chromatoid body.

The mechanisms of polyadenylation are of special interest, because the poly(A) tails on translationally repressed Prm1 and Smcp mRNAs in the spermatid cytoplasm are much longer than the poly(A) tails on active mRNAs, 150 As vs 30 As (Kleene 1989, Yang & Yen 2013). The unusual relationship between poly(A) length and translational activity in spermatids has been highlighted as a significant problem in eukaryotic post-transcriptional gene regulation (Jackson et al. 2010). It is unknown whether the length of poly(A) tails on repressed mRNAs is established by poly(A) lengthening and shortening in the nucleus, chromatoid body, or cytoplasm (Kim & Richter 2006).

Elucidating the mechanisms by which YBX2 and 3T3U YRSs repress translation in round spermatids in mice will have broad significance. The mechanisms of translational regulation of the mouse Prm1 mRNA are at least partly conserved in trout and chicken (Iatrou & Dixon 1977, Rhim et al. 1995). The function of YBX2 as a direct repressor of Prm1 mRNA translation would be solidified by demonstrating that a point mutation in the TCE abrogates YBX2 binding and the strong repression in round spermatids. The analogous studies of the Smcp mRNA should begin with a search for YRSs, because the Smcp 3T3U lacks a recognizable YRS. This search should avoid RNase, because RNase T1 artificially
prevents detection of complexes between YBX2 and the Prm1 TCE (Supplementary Fig. 3). Studies on mouse models should also produce insights into the mechanisms by which mutations in the Ybx2 gene cause abnormal protamine expression and human male infertility (Hammoud et al., 2009).

Supplementary data
This is linked to the online version of the paper at http://dx.doi.org/10.1530/REP-14-0394.

Declaration of interest
The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of this research.

Funding
This research was supported by grant MCB-0642128 to K C Kleene from the National Science Foundation and University of Massachusetts Boston Doctoral Dissertation Grants to D L Cullinane and T A Chowdhury.

Acknowledgements
The authors thank Drs R M Schultz and S Medvedev for the donation of Ybx2+/− mice, the YBX2 antibody, and PCR protocol to distinguish the knockout and WT Ybx2 alleles. They also thank Dr Alexey Veraksa for advice on quantification by mass spectrometry and the use of the confocal microscope.

References
Carter MS & Sarnow P 2000 Distinct mRNAs that encode La autoantigens are differentially expressed and contain internal ribosome entry sites. Journal of Biological Chemistry 275 28301–28307. (doi:10.1074/jbc.M004657200)
Chowdhury TA & Kleene KC 2012 Identification of potential regulatory elements in the 5’ and 3’ UTRs of 12 translationally regulated mRNAs in mammalian spermatids by comparative genomics. Journal of Andrology 33 244–256. (doi:10.2164/jandrool.110.012492)


Received 2 August 2014
First decision 10 September 2014
Accepted 21 October 2014