

FOX-2 SPLICING FACTOR BINDS TO A CONSERVED INTRON MOTIF TO PROMOTE INCLUSION OF PROTEIN 4.1R ALTERNATIVE EXON 16*

Julie L. Ponthier^{‡1}, Christina Schluenzen^{‡1}, Weiguo Chen[§], Robert A. Lersch[‡], Sherry L. Gee[‡], Victor C. Hou[‡], Annie J. Lo[‡], Sarah A. Short[‡], Joel A. Chasis[‡], John C. Winkelmann[§], and John G. Conboy^{‡2}

From the [‡]Lawrence Berkeley National Laboratory, Life Sciences Division, 1 Cyclotron Road, Berkeley, California 94720 and the [§]University of Cincinnati College of Medicine, Division of Hematology-Oncology, Department of Internal Medicine, Vontz Center for Molecular Studies, Cincinnati, Ohio 45267.

Running title: Fox-2 enhances exon 16 splicing in protein 4.1R pre-mRNA

Activation of protein 4.1R exon 16 (E16) inclusion during erythropoiesis represents a physiologically important splicing switch that increases 4.1R affinity for spectrin and actin. Previous studies showed that negative regulation of E16 splicing is mediated by the binding of hnRNP A/B proteins to silencer elements in the exon and that downregulation of hnRNP A/B proteins in erythroblasts leads to activation of E16 inclusion. This paper demonstrates that positive regulation of E16 splicing can be mediated by Fox-2 or Fox-1, two closely related splicing factors that possess identical RNA recognition motifs. SELEX experiments with human Fox-1 revealed highly selective binding to the hexamer UGCAUG. Both Fox-1 and Fox-2 were able to bind the conserved UGCAUG elements in the proximal intron downstream of E16, and both could activate E16 splicing in HeLa cell co-transfection assays in a UGCAUG-dependent manner. Conversely, knockdown of Fox-2 expression, achieved with two different siRNA sequences resulted in decreased E16 splicing. Moreover, immunoblot experiments demonstrate mouse erythroblasts express Fox-2. These findings suggest that Fox-2 is a physiological activator of E16 splicing in differentiating erythroid cells *in vivo*. Recent experiments show that UGCAUG is present in the proximal intron sequence of many tissue-specific alternative exons, and we propose that the Fox family of splicing enhancers plays an important role in alternative splicing switches during differentiation in metazoan organisms.

Alternative splicing of pre-mRNA leads to the synthesis of multiple protein isoforms from a

single gene. It is an important mechanism for regulating gene expression and may be utilized by 40-60% of human genes (1-4). Thus, the estimated 25,000 to 30,000 genes of the human genome can generate a much larger number of proteins. Regulation of alternative splicing occurs in both a tissue- and developmental-specific manner, resulting in alterations in the structure and function of critical proteins. Altered splicing regulation can also be of widespread importance in the etiology of human disease (5-7).

The protein 4.1 gene family serves as an excellent model for investigating the regulation of alternative splicing. The four genes that comprise the family (4.1R, 4.1G, 4.1B, and 4.1N) display a remarkable array of highly regulated, tissue-specific splicing events. These alternative splicing events facilitate expression of distinct isoforms of 4.1 protein in cells of erythroid, epithelial, neural, and muscle origin (8-14); thus, they provide opportunities for understanding the mechanisms that regulate alternative splicing in several different cell types. To date, mechanistic studies have focused predominantly on erythroid cells, in which 4.1R protein is a structural component of the erythrocyte plasma membrane and is important for structural integrity and stability of the membrane skeleton. In differentiating erythroid progenitor cells, a dramatic switch in pre-mRNA splicing results in a physiologically important functional change in the encoded structural protein isoforms (15,16). 4.1R protein in early progenitors, derived from transcripts in which E16 is skipped, exhibits a low affinity for spectrin and actin; in contrast, E16 inclusion in late stage erythroblasts generates a high affinity isoform (17-20).

E16 splicing is influenced by multiple regulatory elements located not only in the exon, but also in the flanking introns, and at the splice sites themselves (21-23). One component of the regulatory machinery in erythroid cells is a stage-specific repression of E16 inclusion, mediated by binding of hnRNP A/B proteins to exonic splicing silencer (ESS) element(s) located within the exon. Repression occurs in early erythroid progenitors, which express high levels of hnRNP A/B proteins, and repression is subsequently relieved by a substantial down-regulation of hnRNP A/B expression in later stage erythroblasts (23). Reduction in the level of splicing inhibitory proteins thus appears to be a critical feature of the physiological E16 splicing switch in erythroid cells.

In this paper, we describe a second mechanism whereby E16 splicing efficiency can be regulated, in this case via the interaction of splicing activator protein(s) with positive regulatory elements in the downstream intron. Our results demonstrate phylogenetic conservation of UGCAUG splicing enhancer motifs downstream of E16, specific binding of these elements to both Fox-1 and Fox-2 splicing factor proteins, and UGCAUG-dependent stimulation of E16 splicing by both of these proteins in functional splicing assays. These findings support the emerging model of Fox protein splicing factors as important components of the cellular machinery that switches on splicing of critical alternative exons at appropriate times during development and differentiation by acting at highly specific UGCAUG intron enhancers (24-27).

MATERIAL AND METHODS

Nomenclature. The splicing factors described in this paper have been reported independently by multiple laboratories in different contexts and with different names. Fox-1 is also known as ataxin 2 binding protein (A2BP1; (28)) or hexaribonucleotide binding protein 1 (HRNBP1; (29)); Fox-2 has been designated as hexaribonucleotide binding protein 2 (HRNBP2; (29)), RBM9 (30), or Fxh (31). Both Fox genes encode multiple isoforms via alternative splicing; e.g., see cDNAs AF094849, AF109106, and AF229057 and recent publications (27,32,33).

Genetic database analysis. Human protein 4.1R exon 16 (May 2004 build, chromosome 1 nucleotides 29207550-29207612) was used as a query sequence with which to retrieve orthologous exons from other genomes, as well as flanking intron sequences, using the BLAT search tools available at www.genome.ucsc.edu/cgi-bin/hgBlat. 4.1G exon 16 is on human chromosome 6 (nt 131242977-131243039); 4.1B exon 16 is on human chromosome 18 (nt 5397436-5397501).

Plasmids. The wild type 4.1R minigene was described earlier (21). In construct 4.1 Δ hex, partial deletion of intron 16 sequences yielded a pre-mRNA that retained 21 nt of proximal intron 16 sequences joined at an XbaI site to the terminal 249 nt of intron 16 and exon 17; all three UGCAUG elements were removed. Derivatives of 4.1 Δ hex were constructed by inserting synthetic oligonucleotides containing two copies of either a wild type (5'-GATCATGCATGAGGGAAAGGTGCATGCAAAGGGAA-3') or a mutated UGCAUG hexamer (5'-GATCATGACTGAGGGAAAGGTGACTGCAAAGGGAA-3').

Mammalian expression clones for mFox-2 α (BC002124) and mFox-2 β (BC027263) were obtained from Research Genetics. The human Fox-1 α cDNA (AF094849) was cloned in one of our laboratories (JCW) and inserted into the expression vector pcDNA3.1-Myc-His(-) between the XbaI and HindIII sites.

siRNAs. Duplex siRNAs obtained from Dharmacon were as follows: Fox-2 siRNA duplex 1: sense strand, 5'-GACAGUAUAUGGUG-CAGUCUU; antisense strand, 5'-PGACUGCA-CCAUAUACUGUCUU. Fox-2 siRNA duplex 2: sense strand, 5'-CGAGAAUAGUGCUGAUG-CAUU; antisense strand, 5'-PUGCAUCAG-CACUAUUCUCGUU.

Splicing analysis. HeLa cells were transfected with plasmid DNAs using Lipofectamine 2000 according to the manufacturer's instructions (Invitrogen). Transfections with siRNA were performed for 24 hrs in 6 well culture dishes using 200 pmol of siRNA. Total RNA was isolated 72 hrs later using RNeasy columns (Qiagen). First strand cDNA was synthesized from the isolated RNA with Superscript II reverse transcriptase (Invitrogen) and an exon 17 3' primer. PCR analysis of spliced 4.1 pre-mRNA was performed using DNA primers located in exon 13 (forward primer: 5'-AGCCATTGCTCAGAGTCAGG-3')

and exon 17 (reverse primer: 5'-GCGAATTCCC GGATT CAGT-3') of the 4.1 minigene.

Immunoblot analysis. HeLa cell protein extracts for immunoblot analysis were prepared by lysis in Novex Tris-Glycine-SDS-buffer (Invitrogen) + DTT and boiled for 10 min. Anti-Fox-2 antiserum (Washington Biotechnology) was raised in rabbits against a synthetic peptide, GVPHTQDYAGQT. This peptide was selected on the basis of its conservation between mouse and human Fox-2 and its divergence from mouse and human Fox-1. The Fox-2 antibody used in these experiments was affinity purified against a synthetic peptide (GenScript Corporation) by using Affi-Gel[®]10 (BioRad) according to the manufacturer's instructions, and was diluted 1:100 in milk blotto/PBS-T solution. For the detection of actin the antibody (Cytoskeleton) was used at 1:5000 dilution. The secondary antibody was goat-anti-rabbit IgG conjugated to HRP (1:5000 dilution). For signal detection, the Western Lightning Chemiluminescence Reagent Kit was employed (Perkin Elmer Life Sciences).

Mouse erythroblasts were obtained from the spleens of mice infected with the Friend virus, anemia-inducing strain (FVA), and cultured as previously described (34,35). Cells differentiate from proerythroblasts to enucleated reticulocytes over ~44-48 hrs in culture. Protein was extracted from the cultured cells and analyzed by immunoblot analysis.

SELEX analysis of Fox-1 binding specificity. For systematic evolution of ligands by exponential enrichment (SELEX) (36), 500 µg of hFox-1 α -myc-his protein, isolated from a baculovirus/SF9-expression system, was coupled to 1 ml NHS-activated Sepharose 4B beads (Pharmacia). The oligonucleotides used for SELEX were identical or similar to those used previously (37,38):

SELN40: 5'-GGCACTATTTATATCAACTAGAACTACTGGATCCG(N)₄₀TTGGTA-CCCAATTCGCCCTATAGTGAGTCGTATTA-3' (N₄₀ representing 40 nucleotides of random sequence)

SELF: 5'-TAATACGACTCACTATAGGGCGAATTGGGTACCAA-3' (for PCR)

SELREV1: 5'-GGCACTATTTATATCAAC-3' (for reverse transcription)

SELREV2: 5'-GGCACTATTTATATCAACTAGAACTACTGGATCCG-3' (for PCR). The SELEX was performed as previously described with some modification (39). 100 µg of SELF and 300 µg of SELN40 were annealed as the template for *in vitro* transcription using RiboProbe[®] *in vitro* Transcription System (Promega). The synthesized RNA was extracted with phenol/chloroform, precipitated, and dissolved in Binding Buffer (0.1 M NaCl, 50 mM Tris-HCl, pH 8.0, 14.4 mM β -mercaptoethanol) containing 4mg/ml tRNA. Approximately 1 mg of RNA was loaded onto the 1 ml hFox1 α -Sepharose 4B column equilibrated with Binding Buffer, mixed and mildly rotated at room temperature for 10 min. The column was washed with 10 ml of Binding Buffer and eluted with 10 ml of 0.1 M NaCl, MTPBS (16 mM Na₂HPO₄, 4 mM NaH₂PO₄, pH 7.3) followed by 10 ml of 1.0 M NaCl, MTPBS. The first 10 fractions (0.5 ml each) eluted by 1 M NaCl, MTPBS were collected, and the RNA was precipitated and dissolved in DEPC-treated H₂O. One fifth of the RNA was reverse transcribed and amplified (22 cycles of 94 °C 45 sec, 65 °C 45 sec and 72 °C 45 sec). The amplified PCR product was resolved by 2% agarose gel electrophoresis. Fractions 2, 3 and 4 eluted with 1 M NaCl, MTPBS were pooled as templates for the next round of SELEX. Beginning with the fourth round, the concentration of NaCl was increased to 0.4 M in both the Binding Buffer and the first elution buffer. A total of seven rounds of SELEX were performed. The PCR product of the final round was digested with KpnI and BamHI and cloned to pBluescript vector. Individual clones were selected and sequenced.

In vitro synthesis of hFox-1 and mFox-2 proteins. Human Fox-1 α cDNA from clone AF094849 was amplified using primers hBP1-S (5'-CACCAGCATGCTGGCGTCTCAAGGAGTTCTC-3') and hBP1-AS (5'-TTAGTATGGAGCAAACGGTTGTATCC-3') and cloned into the bacterial expression vector pEXP1-DEST (Invitrogen). The resulting hFox-1 construct, tagged at the N-terminus with the Express epitope and a six-histidine epitope, was synthesized by *in vitro* transcription/translation using Expressway Plus (Invitrogen) according to the manufacturer's instructions. Mouse Fox-2 β cDNA was amplified from clone BC027263 using primers mBP2-S (5'-

CACCCGGATGGCGGAAGGCGGCCAGGCGC
A-3' and mBP2-AS (5'-
CTAGTAGGGGGCAAATCGGCTGTAGCC-3').
and cloned into pEXP1-DEST for *in vitro*
expression.

*Binding of recombinant Fox proteins to
biotinylated RNAs containing UGCAUG.* 40-mer
RNAs representing nt 79-118 of intron 16
sequence were purchased from Dharmacon. Wild
type RNA sequence: 5' biotin-
GCCCUUGGGUUUGCAUGCCACUGCAUGA
GAGACGUUUmAmG-3'; mutant sequence: 5'
biotin-
GCCCUUGGGUUUGacUGCCACUGacUGAGA
GACGUUUmAmG; (hexamer motifs are
underlined and mutations represented in lower
case). Binding experiments were performed
essentially as previously described (23). Briefly,
recombinant hFox-1 or mFox-2 or hnRNP A1 was
incubated with 100 pmol of biotinylated-RNA in
splicing buffer for 30 min at 4 °C; RNA and
associated proteins were then removed by binding
to streptavidin-conjugated magnetic beads (Dyna) according to the manufacturer's instructions.

RESULTS

*Binding specificity of mammalian Fox splicing
factor proteins for the hexamer UGCAUG.* The
consensus binding site of the zebrafish splicing
factor, zFox-1, was defined as the pentamer
GCAUG (24). We independently cloned human
Fox-1 (hFox-1) cDNA and examined its binding
specificity via SELEX. As shown in Figure 1, all
of the winner sequences binding to hFox-1
contained at least one copy of the same pentamer
recognized by zFox-1, and four contained two
copies (sequence #6, 10, 26, 45). However, closer
examination revealed that the great majority
(45/47) of winner sequences for the human protein
contained the hexamer UGCAUG, with an
additional preference for U in the seventh position,
i.e., UGCAUGU. These results are consistent with
the recent structural analysis of a Fox-1-UGCAUGU
complex (40). This extended consensus binding
site for hFox-1 may reflect evolutionary changes
in the protein from fish to mammals.
Alternatively, the minor difference in binding
specificity may be due to the fact that the reported
zebrafish Fox-1 is not the best ortholog of human
Fox-1; i.e., a closer homolog with higher identity

to the human Fox-1 RRM is encoded in the
zebrafish genome (our unpublished observations).
The powerful selection for winners with multiple
UGCAUG elements may reflect cooperative binding
of hFox-1 to regions with multiple binding sites.
Notably, several of the natural occurrences of
intronic UGCAUG or GCAUG splicing elements,
including the 4.1 gene reported here, contain
multiple elements (22,32,41-44).

Figure 2 shows the domain structure of
mammalian Fox-1 and two closely related Fox-2/
RBM9/HRNBP2 protein isoforms (hereafter called
Fox-2) that are used for functional studies in this
paper. The Fox-2 proteins possess different C-
terminal domains and are derived from the same
gene via alternative splicing, while Fox-1 is
encoded by a separate gene. A C-terminal isoform
of Fox-1 is also encoded by alternative splicing,
but was not studied here. All of these mammalian
Fox proteins are 100% identical in the RRM
domain. Outside of the RRM domain, the proteins
exhibit a reduced but still highly significant
homology, except for the aforementioned
alternative C-terminus. The functional splicing
experiments described below demonstrate that
these novel Fox proteins can promote exon
inclusion in a UGCAUG-dependent manner.

*UGCAUG elements function as intron splicing
enhancers for protein 4.1R exon 16.* E16 in the
4.1R gene undergoes a splicing switch during
erythropoiesis, being skipped in early progenitor
cells and efficiently included at later stages
(Figure 3, top). Three repeats of the UGCAUG
hexamer are located downstream of protein 4.1R
E16 in the human and mouse genes, and mutations
of these elements have been reported to decrease
E16 inclusion (22). As shown in Figure 3, bottom
panel, the presence of UGCAUG elements has been
highly conserved in evolution, as 2-3 repeats are
also located in the E16 proximal intronic regions
of the rat, opossum, chicken, frog, and tetraodon
(fish) 4.1R genes. In addition, two closely
paralogous genes (4.1G and 4.1B) also possess
UGCAUG repeats near alternative E16, although the
surrounding intron sequences are otherwise quite
divergent. These observations support a functional
role for UGCAUG in regulating E16 splicing.

In order to test whether the conserved UGCAUG
elements could function as enhancers or silencers
of E16 splicing, several 4.1 pre-mRNA constructs
possessing or lacking intronic hexamers were

compared in functional splicing assays in transfected HeLa cells. As shown in figure 4, the wild type transcript with three hexamers yielded a modest inclusion efficiency for E16 (43%) among its spliced products (lane 1). In contrast, deletion of an intronic region containing all three hexamers resulted in complete loss of E16 inclusion (construct 4.1 Δ hex; lane 2). This result was consistent with a model in which deletion of an intronic enhancer was responsible for reduced E16 splicing efficiency. Additional constructs were therefore designed to test whether the UGCAUG hexamers represented the active enhancer element(s). Splicing of construct 4.1 Δ hex^{wt}, containing UGCAUG elements reinserted at the site of the deletion, rescued splicing efficiency to near wild type levels (Figure 4, lane 3). Negative control construct 4.1 Δ hex^{mut}, containing a two nucleotide mutation in the inserted hexamer elements, did not rescue E16 splicing (Figure 4, lane 4). Additional construct pairs containing wild type or mutant hexamers with different flanking nucleotides were also tested. While the wild type hexamers consistently yielded much higher exon inclusion efficiency than the comparable mutant hexamers, it was also clear that neighboring sequences could influence the magnitude of these effects (results not shown).

Analogous results were obtained in an *in vitro* splicing assay utilizing HeLa cell nuclear extract. E16 splicing efficiency in the wild type pre-mRNA was dramatically reduced by deletion of the hexamer element; insertion of intact UGCAUG elements, but not mutated hexamers, rescued E16 splicing (data not shown). These results indicated that UGCAUG represents an intronic splicing enhancer for E16 splicing.

Fox splicing factors promote E16 inclusion in a UGCAUG-dependent manner. The results above are consistent with a model in which Fox-1 and Fox-2 function as UGCAUG-dependent splicing factors that activate E16 splicing. This hypothesis was tested directly using a HeLa cell transfection strategy, asking whether Fox-1 or Fox-2 expression plasmids could stimulate E16 inclusion in 4.1 pre-mRNAs. Whereas the transfected wild type 4.1R construct alone exhibited only about 35% inclusion (Figure 5A, 4.1^{wt}), co-transfection with a full length Fox-1 or Fox-2 expression plasmid significantly increased E16 inclusion to 80% or to 70%, respectively. Additional

transfection experiments revealed that despite the difference in C-terminal sequences of Fox-2 α and Fox-2 β , no significant difference in their ability to stimulate E16 splicing was observed (data not shown).

To test whether this observed Fox-protein enhancer activity was dependent on the presence of UGCAUG, further transfection experiments were performed. Pre-mRNA constructs containing or lacking the UGCAUG elements were co-transfected with the Fox-1 or Fox-2 expression plasmids. In contrast to the wild type pre-mRNA transfection, neither protein stimulated E16 splicing in pre-mRNA substrates from which the hexamers were deleted (construct 4.1 Δ hex). Replacement of wild type hexamers (construct 4.1 Δ hex^{wt}) but not mutated hexamers (construct 4.1 Δ hex^{mut}) strongly restored the ability of Fox-1 and Fox-2 to stimulate E16 splicing. The low level of splicing stimulated by Fox-2 in construct 4.1 Δ hex^{mut} is not understood and may indicate that splicing regulation is more complex *in vivo*; however, the major enhancement of splicing by Fox-2 is clearly UGCAUG-dependent.

To confirm, that Fox-proteins are indeed expressed even in the constructs that show no E16 splicing, immunoblot analysis was performed (Figure 5B). As expected, in those cells that contain the Fox expression plasmid, a stronger Fox-1 or Fox-2 signal could be detected (indicated by an arrow) in comparison to all other transfection samples. This observation is not due to a larger amount of protein sample loaded on the gel as demonstrated by the actin detection (Figure 5B, lower panels).

Together, these results suggest that both Fox proteins can rescue splicing in a UGCAUG-dependent manner.

To further confirm the correlation of Fox-proteins and E16 splicing, the effect on splicing was investigated when the protein expression was decreased by RNA interference. Therefore HeLa cells were transfected with the 4.1R minigene construct and siRNA duplexes against the Fox-2 gene. As shown in Figure 6, co-transfection with siRNA duplex 1 led to a substantial reduction of E16 splicing in comparison to the wildtype pre-mRNA. This result correlates with a decrease in Fox-2 expression, demonstrated by immunoblot (Figure 6, middle panel). SiRNA duplex 2 caused

a more dramatic effect: about 50% knockdown of Fox-2 protein was achieved leading to a nearly complete loss of E16 inclusion compared to the 4.1R pre-mRNA. Taken together, these data indicate that Fox-proteins are splicing factors that enhance E16 splicing in HeLa cells.

hFox-1 and mFox-2 bind specifically to the UGCAUG elements in 4.1R intron 16. Several observations suggest that splicing activity of Fox-1 and Fox-2 requires binding to UGCAUG elements in 4.1R intron 16: the SELEX identification of UGCAUG as the binding site for hFox-1; the identity of the RRM domain of Fox-1 and Fox-2; and the UGCAUG-dependence of both proteins' enhancer activities. In order to directly demonstrate UGCAUG-dependent binding, we expressed recombinant Fox proteins by *in vitro* transcription and translation and tested their ability to bind wild type or mutated intron 16 sequences. Biotinylated RNAs representing the region 80-118 nt downstream of E16 were employed in a pulldown assay to assess binding. Each RNA contained either two wild type UGCAUG elements or two mutated UGACUG hexamers. As shown in Figure 7, wild type intron 16 sequence pulled down both hFox-1 and mFox-2 (lanes "hex^{wild}"); in contrast, mutant intron 16 RNA bound much lower quantities of either protein when assayed in parallel under identical conditions (lanes "hex^{mut}"). The immunoblot using anti-hnRNP A1 serves as a control to demonstrate the integrity of the mutant oligonucleotide by virtue of its ability to bind hnRNP A1 similarly to the wild type RNA sequence. Notably, the mutant hexamer was defective both in binding to Fox-1/Fox-2 proteins, as well as in its enhancer activity for E16 splicing. This correlation strongly suggests that binding is directly relevant to enhancer activity.

Expression of Fox-2 in erythroid progenitor cells.

If Fox splicing factors play a role in activating the alternative splicing switch of E16 during erythroid differentiation, one would predict that Fox-2 expression levels should increase relative to that of the silencing factor hnRNP A1 (23). This hypothesis was tested using the *in vitro* differentiation system for mouse erythroblasts isolated from the spleen of Friend virus-infected mice. As demonstrated earlier (23), E16 was excluded in RNA prepared from freshly isolated mouse erythroblasts, but included much more efficiently in RNA prepared from erythroblasts

differentiated in the presence of erythropoietin (Figure 8, upper panel). Protein extracts prepared at time points before and after the splicing switch were then immunoblotted with antibodies to both Fox-2 and hnRNP A1. Fox-2 was detected in the early erythroblasts, and levels changed little in mature cells after the splicing switch (Figure 8, middle panel). In contrast, hnRNP A1 expression was significantly reduced after the splicing switch (Figure 8, lower panel). These results reveal a substantial increase in the Fox-2:hnRNP A1 expression ratio, consistent with a physiological role for Fox-2 in regulating alternative splicing in late erythroid differentiation.

DISCUSSION

In this paper we have demonstrated that the Fox family of pre-mRNA splicing regulatory proteins, acting through UGCAUG intronic enhancer element(s), stimulates inclusion of alternative E16 in protein 4.1R pre-mRNA. These results provide new insights into the tissue-specific control of E16 splicing.

According to our working model, E16 splicing is an ordered process that begins with activation of the weak 5' splice site to facilitate removal of the downstream intron (21). Use of this 5' splice site is inhibited directly or indirectly by binding of hnRNP A/B proteins to a well characterized silencer element in E16 (45), a putative second exonic site (our unpublished results), and potentially in the downstream intron as well (Figure 9). Counterbalancing silencer activity are multiple enhancer elements in the purine-rich region of E16 (23), in the downstream conserved region of the exon (46), and in the downstream intron (22). This suggests a multiple mode of regulation through dynamic antagonism (47) among opposing silencer and enhancer activities. Competition between Fox-1/Fox-2 and hnRNP A1 may be one important determinant of E16 splicing efficiency. Indeed, down-regulation of hnRNP A1 expression in late erythropoiesis relieves the splicing silencing activity to allow activation of E16 splicing (23).

E16 is included not only in late stage erythroblasts, but also in muscle and in brain (9,10,13). Therefore, there may be more than one mechanism for cells to integrate input from distinct differentiation signals to make a splicing

decision to include E16. For example, while downregulation of hnRNP A1 is critical for E16 activation in erythroid cells, increased expression of Fox proteins in muscle and/or brain might be sufficient to switch on E16 splicing, independent of changes in cellular hnRNP A1 levels. Experiments exploring splicing factors interactions during muscle, and neural differentiation will be required to fully elucidate the splicing regulatory networks responsible for E16 regulation in various cell types. The observation that Fox-2 is significantly expressed in mouse erythroblasts suggests that it is a likely physiological activator of splicing in differentiating erythroid cells.

We propose that Fox-2 and its homologs play an important role in tissue-specific alternative splicing regulation in metazoan organisms, perhaps by functioning as part of a molecular switch acting at intronic UGCAUG enhancer elements to activate exon inclusion with the appropriate temporal and spatial specificity. This hypothesis is supported by the observations that UGCAUG is a critical regulatory element for proper tissue-specific splicing of calcitonin/CGRP, fibronectin, myosin II heavy chain-B, and c-src pre-mRNAs (41-44,48); that UGCAUG is statistically highly over-represented in the proximal downstream intron of many brain- and muscle-specific exons (25,26,49); that recombinant Fox-1 promotes muscle-specific splicing in a (U)GCAUG-dependent manner (24); and that Fox-2 enhances inclusion of an exon normally activated during erythroid differentiation (this paper). We speculate that the high binding specificity of Fox proteins is advantageous for splicing switch mechanism(s) designed to activate a limited repertoire of splicing events in response to a specific signaling mechanism in a specific cell type. Increased cellular levels of Fox splicing activity could then facilitate splicing switches in selected target transcripts, analogous to the stimulation of 4.1R E16 splicing that occurs in Fox-transfected HeLa cells, without effecting large general changes in splicing of many pre-mRNAs. The reported upregulation of Fox-2/ RBM9 in

androgen-stimulated spinal motor neurons (31) suggests that it would be interesting to examine these cells for potential Fox-2-mediated activation of brain-specific exons reported earlier to possess proximal intronic UGCAUG elements (25). Alternatively, decreased expression of antagonistic splicing inhibitor proteins could also activate splicing of Fox-responsive exons.

Fox proteins have a long evolutionary history, with close homologs in the genomes of *C. elegans* and *Drosophila* in addition to the zebrafish protein reported earlier. Likewise, the UGCAUG binding sites for Fox splicing factors are conserved not only in the proximal intron sequences of 4.1R E16 (Figure 3), but also adjacent to brain-specific exons from several vertebrate orders (25,26). It will be interesting to elucidate the full extent of Fox splicing factor function in the regulation of alternative splicing in various cell types in higher eukaryotes. Many questions remain to be addressed concerning the mechanism of action for this exciting new class of splicing regulators. One might speculate that Fox proteins function as adaptors to recruit spliceosomal machinery to a regulated exon (50), perhaps synergizing with the action of exonic SR proteins, or with other intronic activators (e.g., (51). Another vitally important issue relates to how Fox enhancer activity could be regulated in a tissue-specific manner, especially given the apparent expression of Fox proteins in multiple tissues. Do the same two (or a few) Fox genes mediate splicing switches in the appropriate temporal and spatial patterns during differentiation of various types of neural and muscle cells, as well as erythroid and potentially other cell types? These issues are likely to be as interesting as they are complex. Analogous to several other splicing factors, the Fox proteins exist as a mixture of alternatively spliced isoforms (27,32,33). Future studies will likely reveal that tissue-specific expression of splicing co-activator proteins, and/or tissue-specific expression of functionally distinct Fox protein spliceoforms, is an essential requirement for precise developmental regulation.

REFERENCES

1. Black, D. L. (2000) *Cell* **103**, 367-370.
2. Brett, D., Pospisil, H., Valcarcel, J., Reich, J., and Bork, P. (2002) *Nat Genet* **30**, 29-30
3. Modrek, B., and Lee, C. (2002) *Nat Genet* **30**, 13-19
4. Roberts, G. C., and Smith, C. W. (2002) *Curr Opin Chem Biol* **6**, 375-383
5. Cartegni, L., Chew, S. L., and Krainer, A. R. (2002) *Nat. Rev. Genet.* **3**, 285-298
6. Faustino, N. A., and Cooper, T. A. (2003) *Genes Dev.* **17**, 419-437
7. Caceres, J. F., and Kornblihtt, A. R. (2002) *Trends Genet* **18**, 186-193
8. Parra, M., Gascard, P., Walensky, L. D., Gimm, J. A., Blackshaw, S., Chan, N., Takakuwa, Y., Berger, T., Lee, G., Chasis, J. A., Snyder, S. H., Mohandas, N., and Conboy, J. G. (2000) *J Biol Chem* **275**, 3247-3255
9. Conboy, J. G., Chan, J. Y., Chasis, J. A., Kan, Y. W., and Mohandas, N. (1991) *J Biol Chem* **266**, 8273-8280
10. Huang, J. P., Tang, C. J., Kou, G. H., Marchesi, V. T., Benz, E. J., Jr., and Tang, T. K. (1993) *J Biol Chem* **268**, 3758-3766
11. Baklouti, F., Huang, S. C., Vulliamy, T. J., Delaunay, J., and Benz, E. J., Jr. (1997) *Genomics* **39**, 289-302
12. Schischmanoff, P. O., Yaswen, P., Parra, M. K., Lee, G., Chasis, J. A., Mohandas, N., and Conboy, J. G. (1997) *J Biol Chem* **272**, 10254-10259
13. Kontrogianni-Konstantopoulos, A., Huang, S. C., and Benz, E. J., Jr. (2000) *Mol Biol Cell* **11**, 3805-3817
14. Ramez, M., Blot-Chabaud, M., Cluzeaud, F., Chanan, S., Patterson, M., Walensky, L. D., Marfatia, S., Baines, A. J., Chasis, J. A., Conboy, J. G., Mohandas, N., and Gascard, P. (2003) *Kidney International* **63**, 1321-1337
15. Chasis, J. A., Coulombel, L., Conboy, J., McGee, S., Andrews, K., Kan, Y. W., and Mohandas, N. (1993) *J Clin Invest* **91**, 329-338
16. Baklouti, F., Huang, S. C., Tang, T. K., Delaunay, J., Marchesi, V. T., and Benz, E. J., Jr. (1996) *Blood* **87**, 3934-3941
17. Discher, D., Parra, M., Conboy, J. G., and Mohandas, N. (1993) *J. Biol. Chem.* **268**, 7186-7195
18. Horne, W. C., Huang, S. C., Becker, P. S., Tang, T. K., and Benz, E. J., Jr. (1993) *Blood* **82**, 2558-2563
19. Schischmanoff, P. O., Winardi, R., Discher, D. E., Parra, M. K., Bicknese, S. E., Witkowska, H. E., Conboy, J. G., and Mohandas, N. (1995) *J Biol Chem* **270**, 21243-21250
20. Gimm, J. A., An, X., Nunomura, W., and Mohandas, N. (2002) *Biochemistry* **41**, 7275-7282
21. Gee, S. L., Aoyagi, K., Lersch, R., Hou, V., Wu, M., and Conboy, J. G. (2000) *Blood* **95**, 692-699
22. Deguillien, M., Huang, S. C., Moriniere, M., Dreumont, N., Benz, E. J., Jr., and Baklouti, F. (2001) *Blood* **98**, 3809-3816.
23. Hou, V. C., Lersch, R., Gee, S. L., Ponthier, J. L., Lo, A. J., Wu, M., Turck, C. W., Koury, M., Krainer, A. R., Mayeda, A., and Conboy, J. G. (2002) *EMBO J* **21**, 6195-6204.
24. Jin, Y., Suzuki, H., Maegawa, S., Endo, H., Sugano, S., Hashimoto, K., Yasuda, K., and Inoue, K. (2003) *Embo J* **22**, 905-912
25. Brudno, M., Gelfand, M. S., Spengler, S., Zorn, M., Dubchak, I., and Conboy, J. G. (2001) *Nucleic Acids Res* **29**, 2338-2348.
26. Minovitsky, S., Gee, S. L., Schokrpur, S., Dubchak, I., and Conboy, J. G. (2005) *Nucleic Acids Res* **33**, 714-724
27. Baraniak, A. P., Chen, J. R., and Garcia-Blanco, M. A. (2006) *Mol Cell Biol* **26**, 1209-1222
28. Shibata, H., Huynh, D. P., and Pulst, S. M. (2000) *Hum Mol Genet* **9**, 1303-1313
29. Winkelmann, J. C., Chen, W.-G., Chu, Z.-L., and Blough, R. I. (2000) *Blood* **96**, Suppl, 592a
30. Norris, J. D., Fan, D., Sherk, A., and McDonnell, D. P. (2002) *Mol Endocrinol* **16**, 459-468

31. Lieberman, A. P., Friedlich, D. L., Harmison, G., Howell, B. W., Jordan, C. L., Breedlove, S. M., and Fischbeck, K. H. (2001) *Biochem Biophys Res Commun* **282**, 499-506
32. Nakahata, S., and Kawamoto, S. (2005) *Nucleic Acids Res* **33**, 2078-2089
33. Underwood, J. G., Boutz, P. L., Dougherty, J. D., Stoilov, P., and Black, D. L. (2005) *Mol Cell Biol* **25**, 10005-10016
34. Koury, M. J., Sawyer, S. T., and Bondurant, M. C. (1984) *J Cell Physiol* **121**, 526-532
35. Lee, G., Spring, F. A., Parsons, S. F., Mankelov, T. J., Peters, L. L., Koury, M. J., Mohandas, N., Anstee, D. J., and Chasis, J. A. (2003) *Blood* **101**, 1790-1797
36. Tuerk, C., and Gold, L. (1990) *Science* **249**, 505-510
37. Tacke, R., and Manley, J. L. (1995) *EMBO J* **14**, 3540-3551
38. Tsai, D. E., Harper, D. S., and Keene, J. D. (1991) *Nucleic Acids Res* **19**, 4931-4936
39. Wang, J., Dong, Z., and Bell, L. R. (1997) *J Biol Chem* **272**, 22227-22235
40. Auweter, S. D., Fasan, R., Reymond, L., Underwood, J. G., Black, D. L., Pitsch, S., and Allain, F. H. (2006) *EMBO J* **25**, 163-173
41. Huh, G. S., and Hynes, R. O. (1994) *Genes Dev.* **8**, 1561-1574
42. Kawamoto, S. (1996) *J. Biol. Chem.* **271**, 17613-17616
43. Hedjran, F., Yeakley, J. M., Huh, G. S., Hynes, R. O., and Rosenfeld, M. G. (1997) *Proc Natl Acad Sci USA* **94**, 12343-12347
44. Lim, L. P., and Sharp, P. A. (1998) *Mol Cell Biol* **18**, 3900-3906
45. Hou, V. C., and Conboy, J. G. (2001) *Curr Opin Hematol* **8**, 74-79.
46. Yang, G., Huang, S. C., Wu, J. Y., and Benz, E. J., Jr. (2005) *Blood* **105**, 2146-2153
47. Charlet, B. N., Logan, P., Singh, G., and Cooper, T. A. (2002) *Mol Cell* **9**, 649-658.
48. Modafferi, E. F., and Black, D. L. (1997) *Mol. Cell Biol.* **17**, 6537-6545
49. Yeo, G. W., Van Nostrand, E., Holste, D., Poggio, T., and Burge, C. B. (2005) *Proc Natl Acad Sci USA* **102**, 2850-2855
50. Guo, N., and Kawamoto, S. (2000) *J Biol Chem* **275**, 33641-33649
51. Forch, P., Puig, O., Martinez, C., Seraphin, B., and Valcarcel, J. (2002) *EMBO J* **21**, 6882-6892

FOOTNOTES

*This work was supported by National Institutes of Health (NIH) grants HL45182 and DK32094, by a research fellowship from the German Research Foundation (DFG) for CS, and by the Director, Office of Biological and Environmental Research, US Department of Energy under contract DE-AC03-76SF00098.

¹These authors have contributed equally to this work.

²Address correspondence to: John G. Conboy, Lawrence Berkeley National Laboratory, Life Sciences Division, 1 Cyclotron Road, Mailstop 74-157, Berkeley, California, 94720, Tel: (510) 486-6973 Fax: (510) 486-6746, E-mail: jgconboy@lbl.gov

FIGURE LEGENDS

Fig. 1. Binding specificity of human Fox-1 for the hexamer UGCAUG. Shown are the winner sequences identified in an *in vitro* selection experiment performed with recombinant hFox-1. The sequences are aligned with respect to the hexamer UGCAUG, shared by 45/47 selected sequences in this experiment (boxed). The remaining two winner sequences (H1-51 and H1-58) possessed a very similar sequence, AGCAUG. Also boxed are the second UGCAUG elements found in four sequences.

Fig. 2. Domain structure of Fox-1 and Fox-2. Shown are three Fox proteins used in this study, including one isoform of hFox-1 and two isoforms of mFox-2. RNA binding domain of each protein is indicated (solid black). Numbers represent the percent identity in amino acid sequence among the three proteins. mFox-2 β possesses a distinct C-terminal domain (solid gray) due to alternative splicing.

Fig. 3. Conserved hexamer elements flanking protein 4.1 exon 16. Upper panel shows a portion of the 4.1R gene from exon 13 to exon 17, and illustrates the splicing switch that occurs during erythroid differentiation. 4.1R mRNA spliced in early progenitors skips E16, and the resulting protein has a low affinity spectrin-actin binding (SAB) domain. Late erythroblasts include E16 efficiently and synthesize 4.1R protein with a high affinity for binding to spectrin and actin. Exons 14 and 15 are skipped in erythroid cells but included in selected nonerythroid cells. Lower panel shows the presence of conserved UGCAUG elements (circles) in intron 16 of 4.1R, 4.1G and 4.1B genes of the indicated species.

Fig. 4. UGCAUG is an intronic enhancer of 4.1R exon 16 splicing. Upper panel shows a model of the 4.1R minigene and 4.1R minigene mutants. RNA 4.1 Δ hex lacks all three UGCAUG elements due to a 186 nt deletion, whereas RNA 4.1 Δ hex^{wt} has four copies of the UGCAUG element reinserted, and RNA 4.1 Δ hex^{mut} has four copies of a mutated element UGACUG reinserted into the deletion construct. Lower panel shows the splicing phenotype of each pre-mRNA. 4.1R minigenes were transfected into HeLa cells and the resulting spliced mRNAs were analyzed by RT-PCR using primers selective for the transfected sequence. Analysis of PCR products by electrophoresis shows a larger product (including E16) in the upper band and a smaller product (excluding E16) in the lower band.

Fig. 5. Fox-1 and Fox-2 are UGCAUG-dependent splicing enhancers for exon 16 splicing. *A*, UGCAUG-dependent Fox protein enhancer activity. Shown is the splicing efficiency of E16 in four pre-mRNA substrates that contain (4.1^{wt} and 4.1 Δ hex^{wt}) or lack (4.1 Δ hex and 4.1 Δ hex^{mut}) UGCAUG splicing enhancer elements. For each construct, six transfections were performed: duplicate transfections without added enhancer; two with co-transfection of Fox-1, and two with co-transfection of Fox-2. Only constructs with UGCAUG enhancers exhibited strong stimulation of E16 splicing in response to co-transfected Fox proteins. Numbers below each lane indicate the percent inclusion of E16. *B*, Expression of Fox-1, Fox-2 and actin verified by immunoblot analysis. The samples described in 5(A) were examined with regard to Fox-1 and Fox-2 protein expression. The transfected Fox-1 with ~50 kDa and Fox-2 with ~52 kDa are indicated by an arrow. Actin serves as indicator for the amount of protein on the gel.

Fig. 6. SiRNA treatment confirms the correlation of Fox-2 and exon 16 splicing. HeLa cells were transfected with the 4.1R minigene alone or in combination with Fox-2 siRNA duplex 1 or duplex 2. The upper panel shows the splicing efficiency of E16. Number below each lane indicates the percent inclusion of E16. In the lower panel immunoblots depict the expression of Fox-2 and actin.

Fig. 7. Binding of Fox-1 and Fox-2 to UGCAUG elements in intron 16. Pull-down experiment with *in vitro*-synthesized Fox-1 (left), Fox-2 (middle) or hnRNP A1 (right) proteins using biotinylated wild type RNA containing two wild type UGCAUG elements or two mutated UGACUG hexamers. Proteins bound to the RNA probes immunoblotted with anti-Fox-1 (left) or anti-Fox-2 (middle) or anti-hnRNP A1 (right).

Fig. 8. Fox-2:hnRNP A1 ratio in mouse erythroid differentiation. Erythroblasts from mice treated with the anemia-inducing strain of Friend virus (FVA cells) were cultivated *in vitro* in the presence of erythropoietin. RT-PCR analysis of E16 splicing patterns in erythroblasts cultured for 24 and 45h (upper panel) and immunoblot analysis detecting Fox-2 (~48 kDa) and hnRNP A1 (~33 kDa) proteins before and after the splicing switch (lower panels). Actin serves as control for the amount of protein loaded on the gel.

Fig. 9. Regulatory elements and splicing factors that control exon 16 alternative splicing. Splicing enhancer elements in E16 (ESE) and the downstream intron (ISE) are proposed to stimulate splicing at the 5' splice site. Further, each enhancer is proposed to be antagonized by adjacent or overlapping silencer elements, including a known hnRNP A1-binding silencer element in the exon (ESS) and a putative silencer in the intron (ISS). Antagonism between the exonic elements has been reported earlier (20). A second putative silencer element may be located in the 5' region of the exon (unpublished results). Three UGCAUG repeats are located within 140 nt downstream of E16.

Figure 1

H1-2		CAUAU	UGCAUG	UUGUGUUGGGCAAAAUUCCUAACUU
H1-3		GUA	UGCAUG	AACUGGUCUUACCUGAAUAGCCUAGUAACUA
H1-4			UGCAUG	AUGGUUGCCGUUAAUUUGAUUUUAACAGCAGGCA
H1-5		CAA	UGCAUG	AUGUUGGUUUCCAACGAUCUCCGAGGACGAU
H1-6		UACUGUUG	UGCAUG	AGGUUAGGUGCCA UGCAUG AUAAUUCG
H1-7		UGUUCACGGA	UGCAUG	GGCGUGGAUUGAGUAUAGACAAUU
H1-8		GUAGGA	UGCAUG	AGCUACUUGGAGGGUUAGCGAGCCAUUG
H1-9		UCAGA	UGCAUG	GGUGAUAGGCCUAUUGCCUUUAACCGCC
H1-10			UGCAUG	GUGGCACCUAUG UGCAUG AUUCAUUCGCUUGGUGU
H1-11	AAUUCUUUUUAAACUUGACCGCAGGUG		UGCAUG	UGC UUG
H1-12		CAGCAGU	UGCAUG	AGUGCUGUGCCCAAAUGCUUAGAUUA
H1-15			UGCAUG	AUGGUCGUAUAACGUUCUAACUUUACCACUA
H1-16		UGAA	UGCAUG	UGACCUUGUAUCCUGUCCGUUCCUAGACU
H1-17		CCGUGAGCUGU	UGCAUG	UGAGUUCUUGGUGGGACCAAUUC
H1-18		CUGAAUUGU	UGCAUG	AU
H1-19		UAGCACAUAUCGGGGGGCAAGU	UGCAUG	UGUUGUCUCU
H1-20		UAGAUGCCCGGGAUCUGCUUUGAGU	UGCAUG	UGUCAAUUG
H1-22	UGUCACCUAGAUAUAUCUGCUGCUUAUCGGU		UGCAUG	UG
H1-23	UGUCCCUUUUAGUGCUGGAGUACGAAGU		UGCAUG	UGUU
H1-24		UUA	UGCAUG	AAAUUGGUUCUAAAUCUUUUAUUUUGGAACUG
H1-25		UCGGG	UGCAUG	UGCGGUUGGUGCUAAUUUGAUGCUCUUGG
H1-26		CCCGAUUAAGU	UGCAUG	UUUAUAAGUCGGGAAA UGCAUG U
H1-27	CCAAAGGGAAUUCUAGCCGGUUUGGU		UGCAUG	UGCAAAC
H1-28			UGCAUG	AUGGUGCACCCGAUUGGCUUAAAUGUGGAGCG
H1-29		GU	UGCAUG	UGUUGGAUCAUUCGUCGCGGGUAAAAACUGU
H1-30		UAAAA	UGCAUG	CUUUA
H1-31		UAUGUACGAGCGUGGGG	UGCAUG	UGCUCGCAGUAACGU
H1-32		GGA	UGCAUG	GGCUUGGUACGUUACAUAUGAUGCGCUUAGA
H1-33		AGU	UGCAUG	UGUUUGGUCCGUUAAUUCGUCGGGUUACGUC
H1-34		UAU	UGCAUG	UUAUCGGUAUCCUUCUGAUUUGGGACGAU
H1-36		UUGAUGCCAACGCCGUGACCGGU	UGCAUG	UUGCAGGACA
H1-37		UUAUCACUACUCGGACCGGGAGGGU	UGCAUG	UGC UUC
H1-39		AUCC	UGCAUG	UAAAGCCUAACUGAAUGUGCUAACUAGGGUU
H1-40			UGCAUG	GUGGUACUACUUCGUUUUAUAAGAGAACAUGA
H1-41		CUA	UGCAUG	AAGUUGGACCUAGAUUUGACGAAUUGCGCU
H1-43		UGAUGUAGG	UGCAUG	UGUACAUAUUGGGAAGGUUCGUU
H1-44			UGCAUG	AUGGCUACCACAUUCGCUAGUGCUAAGAGCUAUA
H1-45			UGCAUG	CUGGUACAUGUUUGCCUAAGGAUCGU UGCAUG UA
H1-47			UGCAUG	GAUGGUCCGACUCGUUUCGCUACUAGAAGAACAUU
H1-49		UGU	UGCAUG	UUGUUGUUACCCUAAAUCAAUUUGGCCUACG
H1-50		CGUAAUUGGAAAAUCCGAGACUGGU	UGCAUG	UGC GGUGU
H1-54		CGUUGU	UGCAUG	UGAAUGUUUGGGCCUAACGACCUAUUGACC
H1-56			UGCAUG	AUGGACACUCGAGUAAGCUCGAAUACACGAACAC
H1-57		UCAAGGGAAUUCUAGCCGGUUUGGU	UGCAUG	UGCAAAC
H1-59		UAGUUUUUCAGUGUGU	UGCAUG	UGACACUGAUUUAACUA
H1-51			UGCAUG	AAGGACCCACGUUUUCCUUCGCCGAAAAGUUC
H1-58	UGUCUCGAGUCAAAUUCUCUUGGAUCGGAGU		UGCAUG	UGUC
Consensus		GU	UGCAUG	U (U/G)

Figure 2

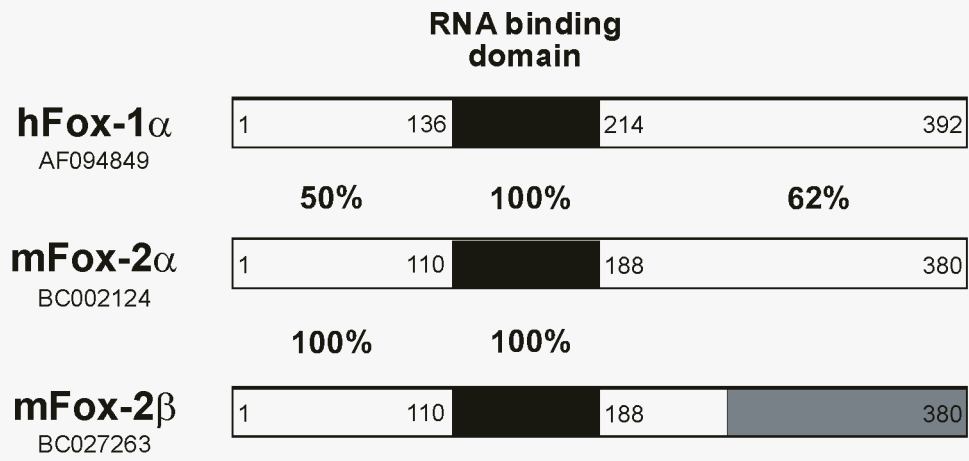
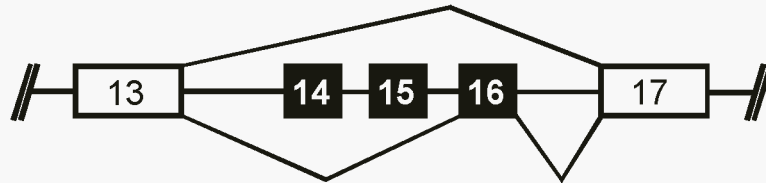


Figure 3

early progenitors
(low affinity binding
to spectrin/actin)



late erythroblasts
(high affinity binding
to spectrin/actin)

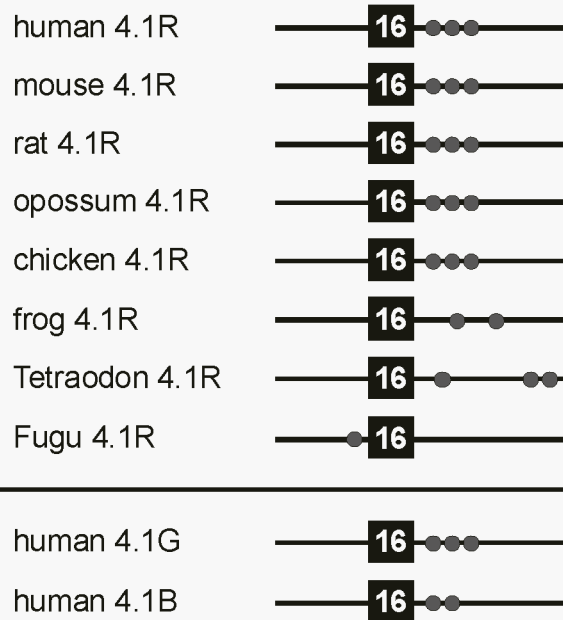


Figure 4



Figure 5A

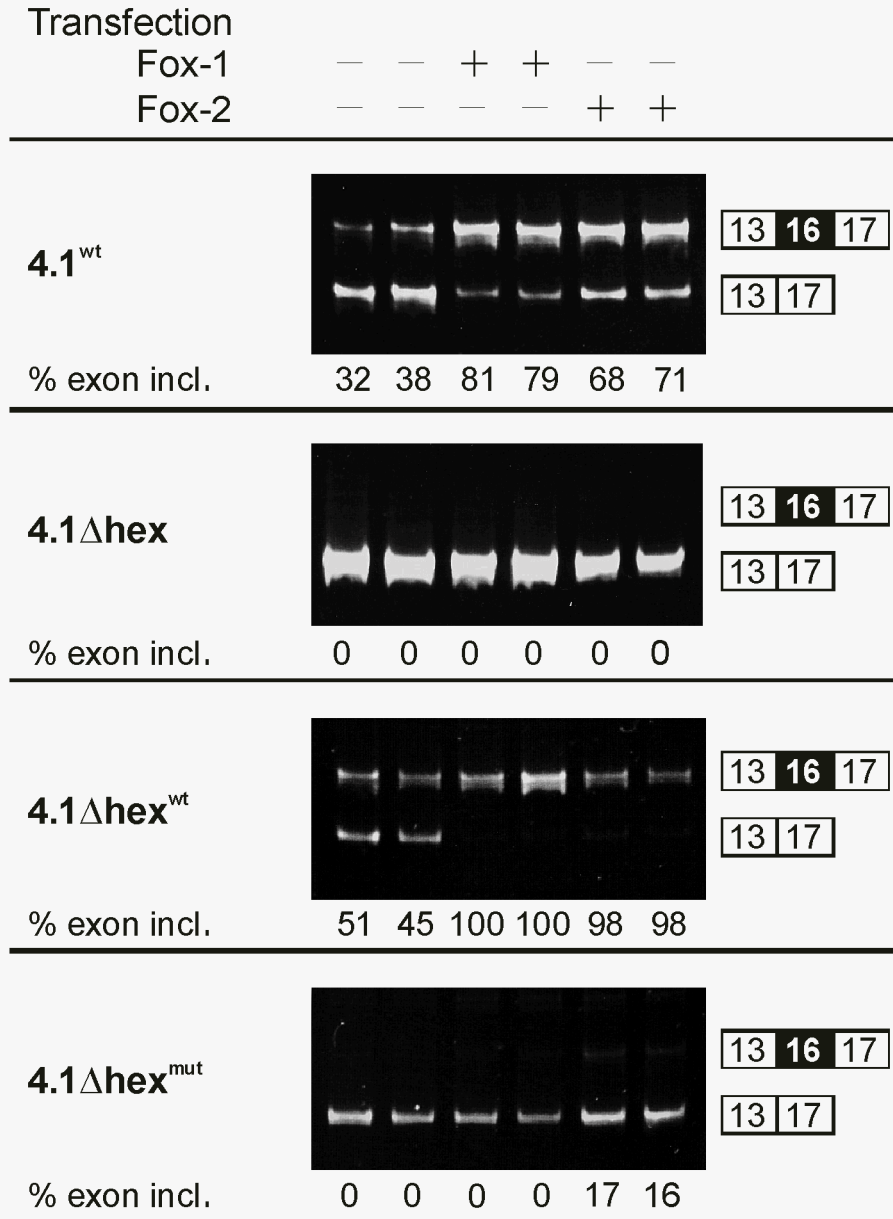


Figure 5B

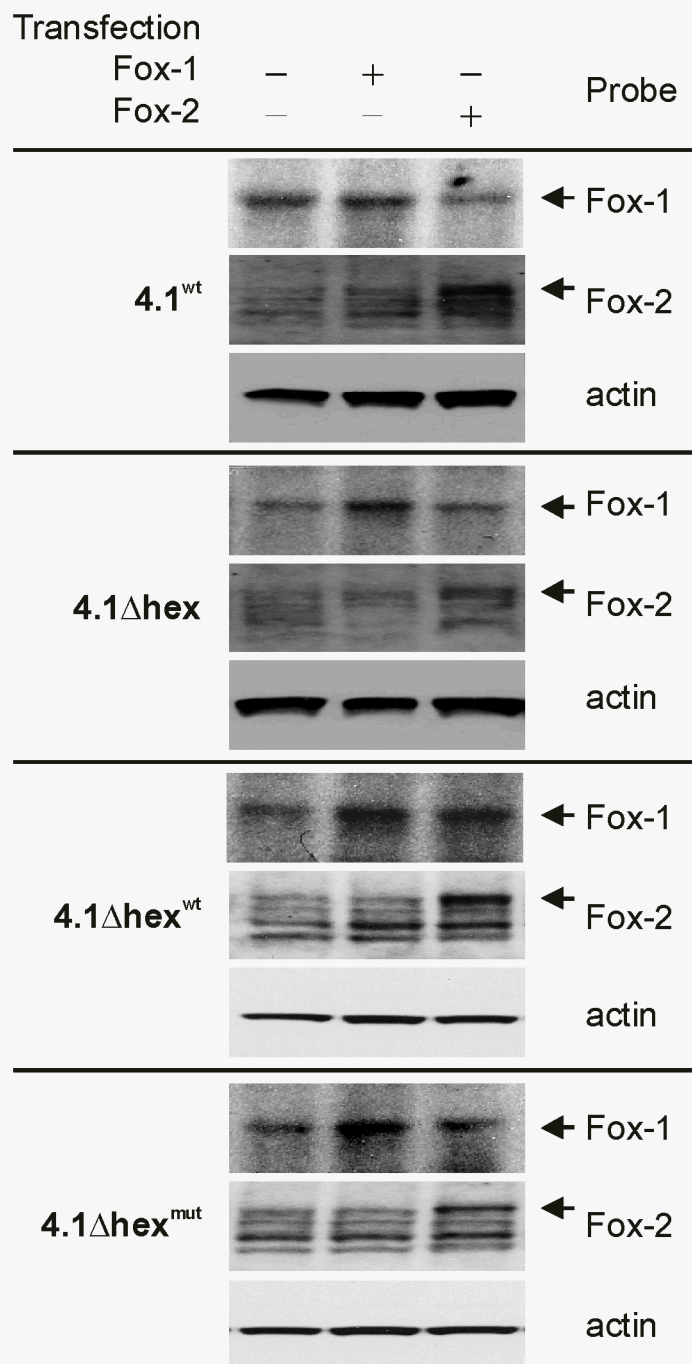


Figure 6

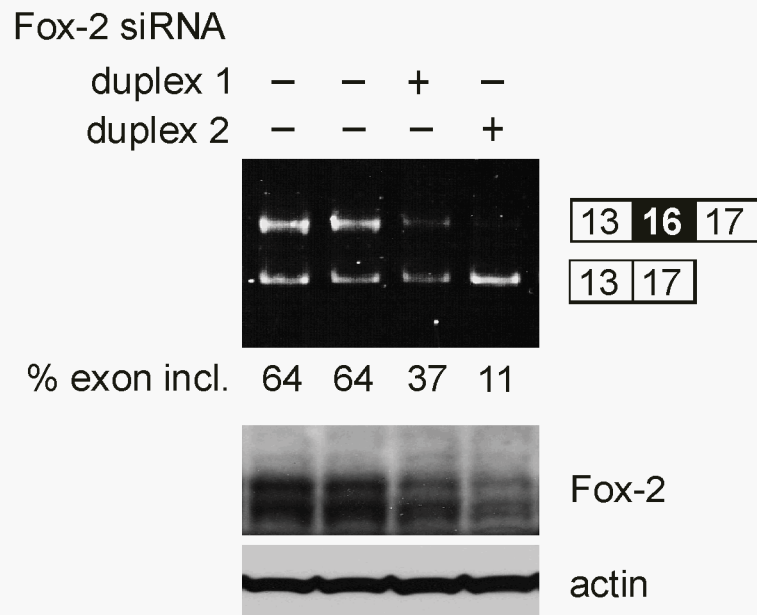


Figure 7

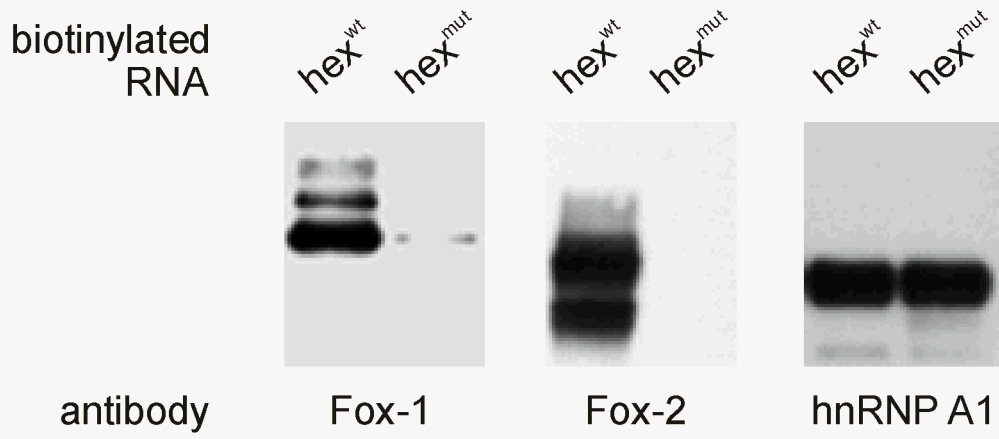


Figure 8

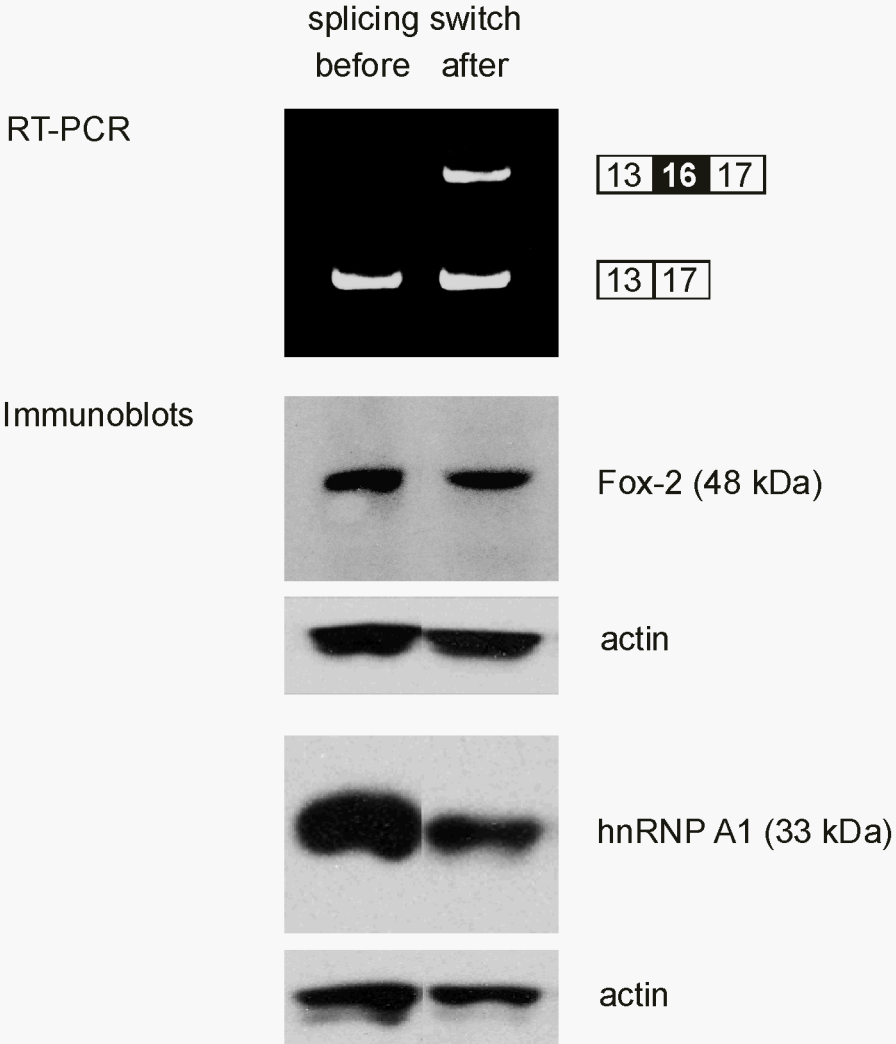


Figure 9

