Functional Pre-mRNA trans-Splicing of Coactivator CoAA and Corepressor RBM4 during Stem/Progenitor Cell Differentiation*

Received for publication, February 2, 2009, and in revised form, April 10, 2009. Published, JBC Papers in Press, May 5, 2009, DOI 10.1074/jbc.M109.006999

Yang S. Brooks, Guanghu Wang, Zheqiong Yang, Kimberly K. Smith, Erhard Bieberich, and Lan Ko

From the Institute of Molecular Medicine and Genetics and Department of Pathology, Medical College of Georgia, Augusta, Georgia 30912

Alternative splicing yields functionally distinctive gene products, and their balance plays critical roles in cell differentiation and development. We have previously shown that tumor-associated enhancer loss in coactivator gene CoAA leads to its altered alternative splicing. Here we identified two intergenic splicing variants, a zinc finger-containing coactivator CoAZ and a non-coding transcript ncCoAZ, between CoAA and its downstream corepressor gene RBM4. During stem/progenitor cell neural differentiation, we found that the switched alternative splicing and trans-splicing between CoAA and RBM4 transcripts result in lineage-specific expression of wild type CoAA, RBM4, and their variants. Stable expression of CoAA, RBM4, or their variants prevents the switch and disrupts the embryonic body formation. In addition, CoAA and RBM4 counter-regulate the target gene Tau at exon 10, and their splicing activities are subjected to the control by each splice variant. Further phylogenetic analysis showed that mammalian CoAA and RBM4 genes share common ancestry with the Drosophila melanogaster gene Lark, which is known to regulate early development and circadian rhythms. Thus, the trans-splicing between CoAA and RBM4 transcripts may represent a required regulation preserved during evolution. Our results demonstrate that a linked splicing control of transcriptional coactivator and corepressor is involved in stem/progenitor cell differentiation. The alternative splicing imbalance of CoAA and RBM4, because of loss of their common enhancer in cancer, may deregulate stem/progenitor cell differentiation.

From the initial discovery of RNA splicing (1, 2) to recent advanced genomic studies (3), a large body of evidence suggests that alternative splicing as an integral part of gene regulation profoundly impacts biological and pathological functions (4). In the human genome, although ~20,000 protein-coding genes exist, more than 90% of multi-exon genes are alternatively spliced (3, 5, 6). Thus, alternative pre-mRNA splicing contributes greatly to proteome diversity without increasing the number of genes (7, 8). Alternative splicing can be cell- and tissue-specific, which is particularly important at the early developmental stages (9). The wild type and variant proteins often exert overlapping but distinct functions whose balance is controlled by the choice of alternative splicing. Aberrant splicing patterns may disrupt the normal functional balance of isoforms, which in turn impairs cell differentiation and induces disease or cancer (10, 11).

Pre-mRNA splicing is a process of joining exons and splicing out introns. In addition to constitutive splicing events, exons can be alternatively spliced. This includes exon skipping, alternative 5’ and 3’ splicing, mutually exclusive exons, intron retention, and alternative transcription start or stop sites (12). The cis-splicing mechanism involves splicing within one pre-mRNA molecule. In contrast, significant evidence supports that trans-splicing can occur at a much lower frequency by joining exons from two independent pre-mRNA molecules through the usage of canonical splicing sites in the absence of DNA recombination (13). Many examples of trans-splicing have been described (14). However, the biological functions of a large amount of trans-splicing events remain to be characterized. Among the products of some alternative splicing and trans-splicing are non-coding RNAs (15, 16). Although non-coding RNAs do not encode a protein product, many of them play important regulatory roles at multiple levels of RNA metabolism. Evidence has also suggested that the majority of the mammalian genome is transcribed into non-coding RNAs, many of which are alternatively spliced products.

Altered splicing patterns are prevalent in diseases and cancers. It has been shown that germ-line sequence variations at splice sites are present in cancer genes such as BRCA1 and APC (17–19). Differentially expressed alternative splice variants are associated with severe clinical outcome in cancer patients (20). Faulty alternative splicing signals impair apoptosis pathways. The high relevance between alternative splicing and cancer is a reflection of the fundamental role of alternative splicing in controlling stem cell or progenitor cell differentiation, as cancer is considered a stem cell disease. Recent studies on human embryonic stem cells and neuronal progenitor cells indicate that a large amount of alternative splicing events normally occur (21). In addition, splicing of a number of exons has been shown to be reprogrammed due to the switch of alternative splicing regulators during neural development (22). Thus, alternative splicing control may play a key role in stem cell and progenitor cell differentiation.
trans-Splicing Variants of CoAA and RBM4

Alternative splicing is coupled with transcriptional regulation which is in part regulated by transcriptional coactivators and corepressors (23, 24). We have previously characterized coactivator CoAA\(^2\) (gene symbol RBM14) in regulating transcription-coupled pre-mRNA alternative splicing (23, 25, 26). The CoAA pre-mRNA is alternatively spliced to yield several alternative splicing variants, one of which is a functional dominant negative variant termed CoAM. In stem cells, CoAA and CoAM regulate early stem cell differentiation through a switch of their alternative splicing at the stage of embryoid body cavitation (27). In human cancers the CoAA gene is amplified with the recurrent deletion of its 5’ upstream regulatory sequence (28). We have shown that deletion of this regulatory sequence leads to defective switching between CoAA and CoAM (27).

In this study we describe the trans-splicing events between CoAA and its downstream RBM4 genes, both of which encode transcription and alternative splicing coregulators that participate in neural stem cell or progenitor cell differentiation. The human RBM4 gene is ~10 kilobases distal to CoAA at chromosome 11q13. RBM4 protein has been previously shown to regulate alternative splicing of the microtubule-associated protein Tau at its exon 10 (29) and to antagonize polypyrimidine tract-binding protein alternative splicing activities (30, 31). RBM4 is a mammalian ortholog of the *Drosophila melanogaster* Lark, which is critical for viability, fertility, development, and circadian rhythms output (32–34). We found that RBM4 represses nuclear receptor-mediated transcription. The CoAA and RBM4 gene transcripts are trans-spliced to produce a novel zinc finger-containing coactivator CoAZ and a non-coding splice variant ncCoAZ. CoAZ and ncCoAZ are derived by 5’ competitive alternative splicing and display switched expression patterns during neural stem cell differentiation. Both CoAZ and ncCoAZ activate transcription and balance coactivator CoAA and corepressor RBM4 splicing activities. In particular, CoAA and CoAZ promote the skipping of Tau exon 10, and RBM4 and ncCoAZ promote the inclusion of Tau exon 10. The imbalance of Tau alternative splicing at its exon 10 is known to be involved in neurodegenerative diseases (35). Thus, the regulation of coactivator CoAA and corepressor RBM4 through alternative trans-splicing variants may have functional importance in normal cell differentiation as well as disease.

**EXPERIMENTAL PROCEDURES**

**Cloning of CoAZ and ncCoAZ**—Total RNA from HeLa cells was isolated using Trizol reagent, treated with DNase I, and reverse-transcribed to cDNA using SuperScript III first-strand synthesis kit (Invitrogen). The presence of CoAZ and ncCoAZ was initially detected by RT-PCR followed by sequence analysis. Full-length cDNAs were subsequently obtained by 3’ and 5’-RACE. As controls, the absence of CoAA transcript in 5’-RACE and the absence of RBM4 transcript in 3’-RACE were confirmed using CoAA or RBM4-specific primers. Primers crossing splicing junctions were used in RACE to ensure specificity. The 5’ and 3’ ends of CoAZ and ncCoAZ transcripts were subjected to sequence analysis. The primers for 5’-RACE were: CoAZ primers, GSP1, tctatcgcagacctctcttg; GSP2, acctgcactgttggctctta; GSP3, tgtgcagcagacgctgtt; ncCoAZ-specific, ttcgaacctgtgattct; The primers for 3’-RACE primers were: CoAZ primers, GSP1, gattcagatctgatagttgcggca; GSP2, tgacgtggtaa-aagcgc; GSP3, cttgtcacaagagttgcagat; ncCoAZ-specific, gataacagactgtggca. Identified cDNAs sequences have been deposited in GenBank with accession numbers as follows: human CoAZ, EU287938; human ncCoAZ, EU287939; mouse CoAZ, EU287940; mouse ncCoAZ, EU287941.

**Construction of RBM4 Minigenes and trans-Splicing Assays**—The CoAA minigenes with alternative splicing capacity have been previously described (27). To facilitate trans-splicing analysis, RBM4 minigenes were constructed under the control of CMV or its native RBM4 promoter (from −1113 to +1) as diagrammed in Fig. 2A. The RBM4 minigenes contain four intact exons, nucleotides 1–71, 1019–1442, 4769–5459, and 7346–7787. Deletions were introduced to the first intron (519–813), the second intron (1623–4558), and the third intron (5728–7065). A minimum 160-bp intronic sequence around each splicing site was left intact to preserve necessary binding sites for splicing factors. A Spel restriction enzyme recognition site was introduced into the third exon of RBM4 (5125–5131) to distinguish minigene transcripts from the endogenous. The splicing capacity of RBM4 minigene cassettes was evaluated by RT-PCR. The trans-spliced transcripts were gel-purified (Qiagen) and SpeI-digested before gel analysis. RT-PCR analyses using primers spanning the CoAA first exon and the RBM4 third exon were performed. Primers used gattcagatctgatagttgcggca and ctaaagctctaagaggtctgacc. PCR products corresponding to the correct size of CoAZ and ncCoAZ were gel-purified and SpeI-digested before gel analysis.

**RT-PCR and Quantitative Real-time PCR Analyses**—Endogenous CoAZ, ncCoAZ, and RBM4 mRNA expression patterns were analyzed using first-strand cDNAs from multiple normal human tissues and cancer cell lines (MTC panels, Clontech). For P19 cells, total RNA was isolated at each differentiation stage using Trizol reagent (Invitrogen), treated with DNase I, and normalized for their concentrations before use. RNA was reverse-transcribed to cDNA using SuperScript III first-strand synthesis system (Invitrogen). Real-time PCR (iCycler, BioRad) was performed using SYBR Green dye in duplicate in a 25-μl reaction. The results were normalized to GAPDH. Primer pairs used were as follows: primers common to endogenous CoAA and CoAM, atggaagtttggcagca, ctaaagctcgcgctgaac; CoAZ-specific, ttcttacacgggatgtgt, cttct-ctgggcgcttgta; ncCoAZ, atcagttgattgctggaaag, aagctttgctcttattcttgctg; CoAZ, tgaagtttgaaagcag, atctttgacactccttgga; RBM4, gctattgctgctgctgactcag, atcttgcgcacactccttgga; total Tau, ccaccaaatccggagaaagcag, gttcgatgtagttgctcca; Tau exon 10, ggcgcataattaataagaagcag, gttcgatgtagttgctcca; Nanog, aggtcgcc-

---

\(^2\) The abbreviations used are: CoAA, coactivator activator; CoAM, coactivator modulator; CoAZ, coactivator with zinc finger; ncCoAZ, non-coding CoAZ; RBM4, RNA binding motif protein 4; ES, embryonic stem; EC, embryonic carcinoma; EB, embryoid body; RBM, RNA recognition motif; GFAP, glial fibrillary acidic protein; RT, reverse transcription; RACE, rapid amplification of cDNA ends; CMV, cytomegalovirus; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; MMTV, murine mammary tumor virus; MAP-2, microtubule-associated protein-2.
trans-Splicing Variants of CoAA and RBM4

Identification of CoAZ and ncCoAZ as trans-Splicing Variants of CoAA and RBM4—We have previously reported that gene amplification and its associated genetic alterations in oncogene CoAA deregulate CoAA pre-mRNA alternative splicing and potentially impact stem cell differentiation (27, 28). At the 5’ boundary of all identified ampiclons, CoAA (RBM14), RBM4, and RBM4B coregulator genes are consecutively aligned and are conserved in mammals. Supported by the NCBI EST database (supplemental Fig. 1) as well as a genome-wide study (14), there are trans-splicing events between CoAA and its downstream RBM4 gene, which is ~10 kilobases distal to CoAA. To identify these trans-splicing variants, total RNA from HeLa cells was extracted. RT-PCR analysis spanning both CoAA and RBM4 followed by sequencing analysis confirmed the presence of two trans-splicing events. As shown in Fig. 1A, CoAA contains three exons, and RBM4 contains four exons. In addition to the alternatively spliced transcripts corresponding to the CoAA and CoAM transcripts we previously reported (25, 27), we identified two additional transcripts, joining the first exon and the RBM4 second or third exons. Further 5’-RACE and 3’-RACE analyses were performed, and the two full-length trans-spliced cDNAs were cloned and designated as
CoAA (coactivator with zinc finger) and ncCoAZ (non-coding CoAZ). Both CoAZ and ncCoAZ share identical transcripts with CoAA at the 5′-untranslated region. The 3′-untranslated region of CoAZ and ncCoAZ is identical to that of the RBM4 transcript, indicating that they utilize the same polyadenylation signals. RACE analyses specific to ncCoAZ utilized primers in the RBM4 second exon (not shown). RACE analyses specific to CoAZ utilized splicing junction primers to enhance specificity (Fig. 1B). The results suggested that CoAZ is an in-frame fusion transcript that encodes a protein of 339 amino acids containing the first RRM domain of CoAA and a CCHC retroviral type zinc finger-containing C terminus (ZnF) is indicated. Deposited accession numbers at NCBI GenBankTM are human CoAZ, EU287932; human ncCoAZ, EU287939; mouse CoAZ, EU287940; and mouse ncCoAZ, EU287941.

CoAZ and ncCoAZ Are Derived from trans-Splicing of Two Separated mRNAs—The donor and acceptor sites for the trans-splicing events in CoAZ and ncCoAZ are separated by ~20 kilobases of sequence in both human and mouse. The possibility that CoAZ and ncCoAZ are produced by trans-splicing events from two independent transcripts rather than by cis-splicing events of a single long transcript is supported by the RT-PCR analyses using primers on different exons of CoAA and RBM4 (supplemental Fig. 2). To further confirm the trans-splicing events, we constructed the minigenes of CoAA and RBM4 and tested them in cells. As shown in Fig. 2A, each minigene contained shortened intron regions to facilitate the subsequent transfection. Both CoAA and RBM4 minigenes were driven by either a CMV promoter or their native promoter sequences. Our rationale was that trans-splicing events would be detected in cells between two independent minigenes if the wild type genes were naturally trans-spliced. Without trans-splicing events, PCR products will not be detected as the DNA templates are on two separated plasmids. The CoAA minigenes were previously described to retain alternative splicing capacities (27). Accordingly, RBM4 minigenes with either the native promoter or CMV promoter were similarly constructed. A Spel restriction site was introduced into the RBM4 third exon to distinguish the minigene transcripts from the endogenous transcripts by restriction digestion. The RBM4 minigenes were first evaluated for their splicing capacities using RT-PCR after transfection of 293 cells. The size of PCR products indicated that all three CoAA and RBM4 splicing events, we constructed the minigenes of CoAA and RBM4 (Fig. 2). To further confirm the trans-splicing events, we constructed the minigenes of CoAA and RBM4 and tested them in cells. As shown in Fig. 2A, each minigene contained shortened intron regions to facilitate the subsequent transfection. Both CoAA and RBM4 minigenes were driven by either a CMV promoter or their native promoter sequences. Our rationale was that trans-splicing events would be detected in cells between two independent minigenes if the wild type genes were naturally trans-spliced. Without trans-splicing events, PCR products will not be detected as the DNA templates are on two separated plasmids. The CoAA minigenes were previously described to retain alternative splicing capacities (27). Accordingly, RBM4 minigenes with either the native promoter or CMV promoter were similarly constructed. A Spel restriction site was introduced into the RBM4 third exon to distinguish the minigene transcripts from the endogenous transcripts by restriction digestion. The RBM4 minigenes were first evaluated for their splicing capacities using RT-PCR after transfection of 293 cells. The size of PCR products indicated that all three RBM4 introns were correctly spliced out, although intron 1 splicing (494 bp) was more efficient (Fig. 2B). The splicing events of RBM4 intron 2 (1117 bp) and intron 3 (884 bp) were present but less efficient, leaving most of the products unspliced. The PCR products in the vector control represented the minigene transcripts from the endogenous spliced transcripts. The gel-purified spliced transcripts were further digested by Spel restriction enzyme.
Our data clearly indicated the presence of digested fragments with the expected sizes that could only be derived from the minigenes (Fig. 2B and supplemental Fig. 3). The endogenous transcripts were not digestible. Because the minigenes retained splicing capacity, we then tested the trans-splicing events by co-transfection of both CoAA and RBM4 minigenes. RT-PCR analyses using primers at the CoAA first exon and RBM4 third exon were performed. The results demonstrated the presence of both CoAZ and ncCoAZ transcripts with digestible products in expected size (Fig. 2C). Our results did not rule out the possibility for trans-splicing between endogenous and minigene transcripts. However, the data nonetheless supported that CoAZ and ncCoAZ can be naturally produced by trans-splicing from two independently transcribed CoAA and RBM4 mRNAs.

CoAZ and ncCoAZ Are Expressed in Normal Human Tissues—To confirm that the trans-splicing products CoAZ and ncCoAZ are not restricted to cancer cell lines we used for cloning but are also present in normal tissues, we investigated the endogenous mRNA distribution of CoAZ and ncCoAZ in multiple human fetal and adult tissues (Fig. 3A). The results suggested that both CoAZ and ncCoAZ mRNAs were widely expressed in normal tissues. The expression of CoAA in these tissues was previously reported (27); RBM4 was also widely expressed. In general, CoAA, RBM4, and ncCoAZ mRNAs are relatively more abundant than CoAZ. In the HeLa cell line, the expression of ncCoAZ was very low but, indeed present, as we cloned the two trans-splicing variants from HeLa cells. Together, CoAA, RBM4, and their splicing variants CoAZ and ncCoAZ are present in normal tissues.
CoAZ Contains RRM and Zinc Finger Domains, and ncCoAZ Does Not Encode a Protein Product—Although trans-splicing phenomena within closely linked genes have been reported, few of the studies detected the presence of a functional chimeric protein (40). To determine whether CoAZ and ncCoAZ transcripts encode protein products in cells, we generated a polyclonal antibody against amino acids 113–339 of CoAZ (β/9251-RRM) that also recognizes the RBM4 zinc finger-containing C terminus. The antibody recognizing the CoAZ RRM domain (β/9251-RRM) can also detect CoAA and its splicing variant CoAM, which was previously described (28). Full-length expression plasmids of each clone with an N-terminal FLAG tag were transfected into 293 cells. The same nuclear extracts from each transfection were concurrently analyzed by Western blots probed with three antibodies against CoAA RRM (β/9251-RRM), RBM4 C terminus (β/9251-ZnF), and FLAG tag (β/9251-FLAG). The results suggested that the CoAZ transcript produced a protein of the predicted size with a CoAA RRM domain and a RBM4 zinc finger-containing C terminus (α-ZnF), and FLAG tag (α-FLAG). The results suggested that the CoAZ transcript produced a protein of the predicted size with a CoAA RRM domain and a RBM4 zinc finger-containing C terminus (Fig. 3B). In 293 cells the endogenous CoAZ was at a much lower level than the endogenous CoAA and RBM4 (Fig. 3B) even though the CoAZ mRNA was detectable (Fig. 3A). Coimmunoprecipitation studies further confirmed that CoAZ protein can be precipitated by both anti-RRM and anti-ZnF antibodies (Fig. 3C). To specifically detect the localization of CoAZ but not RBM4, FLAG-tagged CoAZ was transfected and analyzed by immunofluorescence using anti-FLAG. CoAZ protein was predominantly nuclear in localization (Fig. 3D). In contrast, ncCoAZ did not yield any detectable protein product, which was shown by anti-FLAG Western blot (Fig. 3B). These findings are consistent with our sequence analyses described above that ncCoAZ carries a premature stop codon. Although the ncCoAZ transcript contains a complete coding region of RBM4 whose translation start codon ATG is at the very N terminus of the RBM4 second exon (Fig. 1A), the different sequence in ncCoAZ immediately upstream of the RBM4 start codon may prevent RBM4 protein translation. In this regard, the expression of the ncCoAZ transcript may compete with the expression of RBM4 protein, thus providing a linked control of individual splicing forms through the control of trans-splicing. Therefore, the non-coding transcript ncCoAZ would equally have functional importance as the coding transcript CoAZ.

CoAZ and ncCoAZ Stimulates Nuclear Receptor-mediated Transcription—Because CoAA is a nuclear receptor coactivator, we then investigated the transcriptional activities of CoAZ and ncCoAZ. Each protein domain from CoAA and RBM4 was fused to the Gal4 DNA binding domain and tested on a Gal4-luciferase reporter. Neither the RRM domains nor the zinc finger-containing C-terminal region of RBM4 showed any transcriptional activity when compared with the CoAA activation domain (Fig. 4, A and B). These data indicated that RBM4 and CoAZ do not contain an intrinsic activation domain responsible for stimulation of transcription. We further studied the transcriptional activities of full-length clones in a glucocorticoid receptor-mediated transcription assay using MMTV-luciferase as a reporter and dexamethasone as a hormone ligand.
This system was chosen for its high sensitivity and specificity in measuring transcriptional coregulator activities, although other transcriptional reporter systems could be used to yield similar results. Although CoAA stimulated transcription, RBM4 strongly inhibited transcription. Both CoAZ and ncCoAZ stimulated transcription in a hormone and receptor-dependent manner (Fig. 4, C and D). Their combined effects using cotransfection assays also confirmed that RBM4 potently repressed CoAZ- or ncCoAZ-induced transcriptional activation (Fig. 4E). CoAA contains an activation domain with YXXQ repeats (25). RBM4 lacks an activation domain and possesses a suppressive zinc finger C terminus. Therefore, the transcriptional stimulatory effect of CoAZ may be indirect, through competing corepressor RBM4 with the identical zinc finger in the absence of N-terminal double RRM domains of RBM4 (Fig. 4A). On the other hand, ncCoAZ potently stimulates transcription in the absence of a protein product. This is potentially because of the inhibition of endogenous RBM4 expression as they share splicing sites. Thus, both coactivator variants could inhibit corepressor RBM4 via alternative trans-splicing.

**Endogenous Expression of CoAA, RBM4, and Their Variants during Neural Differentiation**—To study the biological importance of CoAA, RBM4, and their variants, we examined their endogenous expression using immunohistochemistry and immunofluorescence in several physiological contexts. Mouse embryonic stem cells can be induced to form non-adhering cell aggregates called embryoid bodies (EB) that resemble the inner cell mass of embryos. The embryoid bodies reflect an early stage of neural differentiation. Further differentiation yields a mixture of cell types including neurons and glial cells as well as partially differentiated progenitors. At 2 days of embryoid body formation (EB2), both CoAA and RBM4 were highly expressed (Fig. 5A). Although anti-ZnF antibody recognized both RBM4 and CoAZ, RBM4 was far more abundant than CoAZ at this stage (see below). At 4 days of differentiation (EB4), the embryoid bodies formed cavities where apoptotic cells were positive for active caspase-3. The expression of CoAA and RBM4 were reduced in the cavities but were significant in the two primitive ectodermal and endodermal layers which undergo further differentiation (Fig. 5A). At this stage, CoAM, recognized by anti-RRM, was known to be in the embryoid body cavity, and its expression was cytoplasmic (27). Supporting this observation, embryonic mouse brain at gestational stages of E12.5 and E15.5 had high levels of protein expression recognized by anti-ZnF antibody (Fig. 5B). This pattern largely overlapped but was not identical to the CoAA expression pattern we analyzed previously (27). However, both expression levels were high at the stage of E15.5 in the cortical plate, which gives rise to the cerebral cortex. Neurogenesis is at its peak in this stage. The data suggested that CoAA and RBM4 or CoAZ may be differentially involved in neuronal cell differentiation. Further immunofluorescence studies showed that these proteins were indeed expressed in ES cell-derived early stage neural progenitors after 3 days (NP3) and 5 days (NP5) of culture (Fig. 5C), and differ-

**FIGURE 4.** CoAZ and ncCoAZ stimulate transcription and promote CoAA activity. A, domain structures of CoAA, CoAM, CoAZ, and RBM4 proteins are depicted with open boxes as RRM s and filled boxes as the YXXQ or the ZnF domain. CoAM inhibits CoAA via RRM s, and CoAZ inhibits RBM4 via zinc finger domain. B, each depicted CoAA and RBM4 domain was fused to a Gal4 DNA binding domain for testing transcriptional activity. CV-1 cells were transiently transfected with Gal4-fusion proteins and a luciferase reporter containing five Gal4-binding sites. C, full-length CoAA, CoAM, RBM4, CoAZ, or ncCoAZ was tested for transcriptional activity using a MMTV-luciferase reporter in CV-1 cells. Each clone (200 ng) was cotransfected with a MMTV-luciferase reporter (100 ng) and an increasing amount of glucocorticoid receptor (0, 0.1, 1, 10 ng) which binds to the MMTV promoter. Cells were induced by glucocorticoid receptor ligand dexamethasone (100 nm) overnight, and the luciferase activity was measured by Dynex luminometer. D, similar transfections as in C except using 10 ng of glucocorticoid receptor and an increasing amount of dexamethasone (0, 1, 10, 100 nm). E, cells were cotransfected as indicated in the presence or the absence of dexamethasone (100 nm). Luciferase activities shown are the means of triplicate transfections ± S.D.
trans-Splicing Variants of CoAA and RBM4

A

Control  α-CoAA  α-RRM  α-ZnF  α-Active caspas 3

α-RRM  α-CoAA  CoAA  CoAM  CoAZ  ZnF  RBM4  ZnF
α-ZnF

EB2

EB4

B

neocortex  ectoderm  subventricular  intermediate  marginal
surface  zone  zone  zone  zone

α-ZnF

E12.5  E15.5

D

CoAA  RRM  ZnF

Merge
CoAA  RRM  ZnF
GFAP
MAP-2

Brain sections

C

Nestin  GFAP  GFAP

CoAA

RRM

ZnF

NP3 progenitors  NP5 progenitors  Astrocytes

E

MAP-2  DAPI

CoAA

RRM

ZnF

Cortical cells
trans-Splicing Variants of CoAA and RBM4

FIGURE 6. Alternative splicing switch between CoAZ and ncCoAZ during P19 embryonal carcinoma cell differentiation. A, undifferentiated P19 cells (EC) were induced by 500 nM retinoic acid in suspension culture up to 4 days to form embryoid bodies (EB2–EB4). Cells were further differentiated for an additional 15 days in the absence of retinoic acid (D3–D15). Total RNA was isolated and normalized at each stage and analyzed using gene-specific primers by RT-PCR as indicated. GAPDH was control. 15 days in the absence of retinoic acid (D3–D15). Total RNA was isolated and normalized at each stage and analyzed using gene-specific primers by RT-PCR as indicated. GAPDH was control.

Switched Expression of CoAZ and ncCoAZ mRNAs during Stem/Progenitor Cell Differentiation—We have demonstrated in a previous study the presence of switched expression of CoAA and its splice variant CoAM during P19 cell differentiation. We then analyzed the expression patterns of RBM4, CoAZ, and ncCoAZ using this system. P19 is a teratocarcinoma-derived pluripotent stem cell line that can give rise to all three germ layers in mice (41). Retinoic acid induces undifferentiated P19 cells (EC) to form embryoid bodies (EB2–EB4) which continue to differentiate into multiple neural lineages (D3–15). Total RNA was isolated at each time point during differentiation followed by RT-PCR analysis of CoAA, RBM4, and their variants. These PCR products were sequenced to confirm the identities of each variant. Consistent with our previous findings for the switched expression of CoAA and CoAM, the expression of CoAZ gradually increased, whereas the expression of ncCoAZ decreased at a much later differentiation stage (Fig. 6A), which was confirmed by real-time PCR quantitation (Fig. 6B). The data indicated that a switch of CoAZ and ncCoAZ expression has occurred as both utilize the same 5’ alternative splicing site. The expression of RBM4 increased during initial differentiation and decreased at a later stage. Other markers were concurrently analyzed as controls during differentiation. Nanog is known to be decreased during initial stem cell differentiation (42). MAP-2 and GFAP are neuronal and glial cell markers, respectively (Fig. 6A). Of note, there were potential

FIGURE 5. Endogenous expression of CoAA, RBM4, and CoAZ during neural differentiation. A, the diagram at the left indicates antibodies recognizing each variant. ES cell-derived embryoid bodies at 2 days (EB2) and 4 days (EB4) of differentiation stage were paraffin-embedded, sectioned, and stained with affinity-purified anti-CoAA (CoAA only), anti-RRM (CoAA, CoAM, and CoAZ), anti-ZnF (RBM4 and CoAZ), and anti-active caspase-3 (cleaved caspase-3) antibodies and counterstained with hematoxylin. Enlarged views are shown below. B, immunohistochemistry analyses of mouse embryonic brain at gestation stages of E12.5 and E15.5. The sagittal sections were stained with anti-ZnF antibody (1:200). C, immunofluorescence staining of ES cell-derived neural progenitors after 3 days (NP3) or 5 days (NP5) of culture using anti-CoAA, anti-RRM, and anti-ZnF antibodies co-stained with the neural progenitor marker nestin or astrocyte cell marker GFAP (left and middle panels). White arrows indicate negatively stained cells. Double staining of isolated astrocytes from brain of P3 neonates using GAFP (right panel). D, mouse brain sections from P3 neonates were stained with anti-CoAA, anti-RRM, and anti-ZnF antibodies and co-stained with GFAP and MAP-2. E, double staining in MAP-2 positive neurons or MAP-2 negative astrocytes from isolated rat brain cortical cells at E18.5 after 6 days of culture in vitro. Slides are counter-stained with 4’,6-diamidino-2-phenylindole.
correlations between CoAZ and GFAP, which is an astrocyte marker at the later stages of differentiation. In addition, changed levels of CoAZ and RBM4 detected by Western blot were consistent with their mRNA levels (Fig. 6C). A 35-kDa band corresponding to CoAZ was detected by both antibodies. Furthermore, the increase of CoAZ and the decrease of RBM4 were also detected by Western blotting in primary cortical cells of rat brain, which reflects a later stage of neural differentiation (Fig. 6D) (38). These data together suggested that the alternative splicing including trans-splicing events between the CoAA and RBM4 genes are switched during stem cell differentiation, resulting in differential expression of each form in different cell lineages.

CoAA, RBM4, and Their Variants Regulate Alternative Splicing of Tau Exon 10—Alternative splicing control of CoAA and RBM4 genes leads to linked expression of functionally related variants. We then overexpressed each clone in 293 cells and analyzed the mRNA levels of neighboring variants. The results suggested that at the mRNA expression level, CoAA inhibits CoAM and CoAZ inhibits ncCoAZ and vice versa (Fig. 7, A and B). Quantitative PCR analysis confirmed that CoAZ and ncCoAZ had a reciprocal relationship in their expression (Fig. 7C). In addition, ncCoAZ was able to repress endogenous RBM4 expression and to alter the balance of CoAA and CoAM (Fig. 7A). To avoid overlapping with ncCoAZ, RBM4 was analyzed only for its endogenous mRNA. These expression patterns can be clearly explained by their shared alternative splicing sites and potential feedback regulation (Fig. 1A). These results indicated that the expression of CoAA, RBM4, and their variants are linked through alternative splicing and trans-splicing. CoAA is a dominant coactivator, and RBM4 is a dominant corepressor. Each can be competed by its cell-specific splicing

FIGURE 7. CoAA, RBM4, and their variants regulate alternative splicing of Tau exon 10. A, 293 cells were transfected with pcDNA3 vector, CoAA, CoAM, CoAZ, RBM4, or ncCoAZ followed by RT-PCR analyses using gene-specific primers of each transcript as indicated. The detected signals represent both endogenous and overexpressed transcripts except for RBM4. The RBM4 transcript overlaps with the ncCoAZ transcript, and only endogenous RBM4 was analyzed. B, quantitative PCR analysis of mRNA levels of RBM4, CoAZ, and ncCoAZ (200 ng) transfected 293 cells. C, an increasing amount of CoAZ or ncCoAZ (50, 200, 400 ng) was transfected before quantitative PCR. D, the diagram at the left illustrates alternative splicing of Tau exon 10. Primers on exons 9 and 11 detect total Tau mRNA, and primers on exons 10 and 11 detect Tau exon 10 inclusion isoform only. Middle panel, data are shown as real-time PCR quantification of endogenous expression levels of Tau exon 10 inclusion (Tau + Exon 10), total Tau, and their ratio. RT-PCR gel is shown below. The mRNA samples of P19 cell differentiation in this assay were the same preparation as in Fig. 6A. Right panel, Tau alternative splicing analysis in undifferentiated P19 stem cells (EC) overexpressed with CoAA, RBM4, and their variants as indicated. RT-PCR gel is shown below the PCR quantification.
variant CoAM or CoAZ, respectively, through shared functional domains (Fig. 4A). Their overall expression levels and alternative splicing balance are further subjected to regulation by yet another non-coding variant, ncCoAZ, which is efficiently produced by trans-splicing.

To test this model on a physiological target gene, we chose to analyze Tau, as RBM4 has been previously identified by another group through a library screen that binds to Tau mRNA and promotes Tau exon 10 inclusion (29). The microtubule-associated protein Tau is important for neuronal function. Exon 10 inclusion or skipping of Tau produces spliced isoforms whose imbalance is associated with multiple neurodegenerative diseases collectively called “tauopathies” (35). We then tested if CoAA and the variants also regulate Tau exon 10. PCR primers were designed on exon 9 and 11 for total Tau mRNA expression and on exon 10 and 11 for exon inclusion only (Fig. 7D). Consistent with the previous findings that the fetus has more exon skipping and the adult has more exon inclusion (43), RT-PCR and real-time quantification showed that the endogenous Tau exon 10 inclusion gradually decreased and then increased during P19 cell differentiation (Fig. 7D, left panel). The total amount of Tau expression peaked in parallel with MAP-2 (Fig. 6A), suggesting that Tau may be involved in neuronal functions. We then overexpressed CoAA, RBM4, and their variants CoAM, CoAZ, and ncCoAZ in undifferentiated P19 stem cells. The vector was a control representing the endogenous Tau levels. The results indicated that CoAA and CoAZ promoted exon 10 skipping. RBM4, CoAM, and ncCoAZ promoted exon 10 inclusion (Fig. 7D, right panel). In neurons, balanced Tau splicing may be maintained by both CoAA and RBM4, which are present in the cortical plate (Fig. 5B). During gliogenesis, the absence of CoAA and the presence of RBM4 in subpopulations of astrocytes (Fig. 5F) may explain the increased Tau exon 10 inclusion at the late stage of P19 cell differentiation. The data together suggested that CoAA and RBM4 counter-regulate Tau exon 10 splicing. The differential expression of CoAA, RBM4, and their variants potentially contribute to cell-specific expression of Tau isoforms.

Using stably transfected stem cells, we further investigated CoAA, RBM4, and their variants in embryoid body formation. The choice of stable overexpression rather than small interfering RNA (siRNA) knockdown was made because of the lack of sequence specificity for small interfering RNA siRNA in overlapped variants. The constant overexpression by the CMV promoter would disrupt the switched expression patterns required for embryoid body formation. Compared with the vector control, the stable CoAA, CoAM, RBM4, and ncCoAZ cells had defects in formation of embryoid bodies (Fig. 8A). Stable CoAZ cells did not significantly affect the formation of embryoid bodies possibly because of the requirement of CoAZ in a later stage of differentiation. The results for each transfected stable cell were consistently seen in more than two selected stable clones. The vector control showed no detectable difference from untransfected cells in the aspects of embryoid body formation or marker gene expression (not shown). When marker gene expression was analyzed by PCR, the results showed the altered expression levels in both neural lineage markers MAP-2, GFAP, and Tau as well as non-neural lineage markers Nanog and Sox6 (Fig. 8B). In particular, the expression of Tau was elevated before retinoic acid induction in CoAA CoAM, ncCoAZ, and RBM4 stable cells. The exon 10 inclusion of Tau was increased in ncCoAZ and RBM4 stable cells. The data were confirmed by quantitative PCR analysis (Fig. 8B). Although the stable expression did not reflect physiological conditions, the data implicated that the balance of CoAA and RBM4 and their variants is involved in controlling stem/progenitor cell differentiation. Because CoAA and RBM4 are coregulators in multiple tissues, their target genes may include neural lineage genes such as Tau but may not be limited to neural lineage. The imbalance of CoAA and RBM4 activities through the control of their variants can disrupt normal differentiation program.

**DISCUSSION**

In this study we have identified trans-splicing events between the CoAA and RBM4 genes, which yield a novel zinc finger-containing coactivator CoAZ and a non-coding splice variant ncCoAZ. Based on the collective evidence, we propose a model in which the C-terminal activation domain of CoAA stimulates transcription and the C-terminal zinc finger domain of RBM4 represses transcription. RBM4 competes with CoAA through highly homologous double RRM domains at their N termini, which might be responsible for the pre-mRNA interaction during transcription-coupled alternative splicing. CoAM, an alternative splice variant containing only the double RRM of CoAA, has been previously shown to repress CoAA activity. CoAZ contains an intact RBM4 zinc finger but defective double RRM and, thus, competes with RBM4 to indirectly facilitate the CoAA activity. ncCoAZ, although producing no protein product, regulates both CoAA and RBM4 expression via competitive splicing or by participates in potentially unidentified downstream RNA metabolism that impacts both CoAA and RBM4 transcripts. Thus, coactivator CoAA, corepressor RBM4, and their splice variants CoAM, CoAZ, and ncCoAZ have established a linked control because of the presence of alternative splicing and trans-splicing. When they are specifically expressed in different cell lineages, alternative splicing or trans-splicing would constitute a mechanism in regulating the coregulator balance in each cell lineage.

The phenomenon of trans-splicing has been found in a variety of organisms including trypanosomes (44), Caenorhabditis elegans (45), Euglenoids (46), flatworms (47), Drosophila (48), and mammals (49). It is frequently found within a cluster of genes using canonical splice sites and generating translatable mRNA (40). In mammals, the consecutive arrangement of the three genes CoAA (RBM14), RBM4, and RBM4B is highly conserved. All three genes encode RRM-containing proteins and function as alternative splicing regulators. Our study on trans-splicing between CoAA and RBM4 implies that these genes are associated as a gene cluster and share a common enhancer sequence. Expressed sequence tag evidence also supports the presence of splicing events between RBM4 and RBM4B, although at a lower frequency (supplemental Fig. 1). The RBM4B gene encodes an RBM4 isoform (30), and both proteins are mammalian orthologs of the Drosophila protein Lark (32), which was originally identified in Drosophila circadian rhythm regulation and early development (34). We performed a phylo-
genetic analysis using 31 well studied human and *Drosophila* RRM domains including those from CoAA, RBM4, RBM4B, and Lark and found that the two RRM domains in *Drosophila* Lark are most closely related to the two RRM domains in human CoAA (Fig. 9A). Human RBM4 and RBM4B are also closely related to each other. Thus, the Lark gene possibly shares common ancestry with the entire mammalian gene cluster including CoAA, RBM4, and RBM4B (Fig. 9B). CoAA is a coactivator and alternative splicing regulator (23, 25) involved in embryoid body differentiation (27). RBM4 is a corepressor and has been previously identified as alternative splicing regulator that shares nuclear import pathways with SR proteins (30) and antagonizes polypyrimidine tract-binding protein splicing activities (31). The RRM domain of *Drosophila* Lark is involved in maternal function, flight behavior, and developmental control (33). The zinc finger domain of Lark impacts wing morphology and flight behavior. CoAA and RBM4 function in stem or progenitor cells reflects Lark in development. Their functional similarity underscores alternative splicing controls in differentiation and development and further supports that the trans-splicing of CoAA and RBM4 is functionally necessary and, therefore, preserved during evolution.

The CoAA gene may have evolved at a relatively early stage of evolution, whereas RBM4 and RBM4B were derived by gene duplication at a later stage. This prediction is intriguingly consistent with a proposed theory from the studies of the tumor necrosis factor receptor type 2 gene (*p75TNFR*) and many other similar cases, in which alternative splicing is required during gene evolution (12). The theory suggests that newly evolved exons are first introduced by *Alu* repeat-containing transposable elements. Alternative splicing has subsequently evolved to neutralize the potential impact of the creation of a new exon.

**FIGURE 8.** Stable expression of CoAA and RBM4 in P19 cells disrupts embryoid body cavitation. A, light microscopy images of P19 stable cells expressing CoAA, CoAM, CoAZ, ncCoAZ, and RBM4 under the CMV promoter. The uninduced P19 stable cells (EC) and in retinoic acid-induced embryoid bodies at EB4 stages are shown (×400). Vector pcDNA3 was a control. B, RT-PCR analyses of Tau isoforms with (Tau + Exon 10) or without (Tau - Exon 10) exon 10 inclusion in stable P19 cells at differentiation stages of EC, EB2, EB4, and D3. The marker genes used include Nanog, Sox6, MAP-2, and GFAP. GAPDH was a control. Quantitative PCR analysis of total Tau and percentage of Tau exon 10 inclusion are shown below.
required in higher species. The RBM4B isoform possibly has cell-specific expression distinguishable from RBM4 (53). The evolution of the CoAA gene cluster reflects the increasing complexity of splicing regulation because of the increasing cell- and tissue-specificity.

As coregulators, CoAA and RBM4 would have a number of target genes that remain to be identified. Tau has previously been identified as a RBM4 target gene (29), and this study indicates that both CoAA and RBM4 cooperatively regulate Tau alternative splicing. Tau is involved in the intracellular filamentous deposits that define multiple neurodegenerative diseases, including Alzheimer disease, progressive supranuclear palsy, corticobasal degeneration, Pick disease, and the inherited frontotemporal dementia and parkinsonism linked to chromosome 17, collectively known as tauopathies (50, 51). Tau exon 10 encodes a microtubule binding repeat whose inclusion or exclusion yields Tau isoforms with a different number of repeats. The delicately balanced Tau isoforms are critical for neuronal function in learning and memory (52), and splicing mutations of Tau have been found in neurodegenerative diseases (35). The splicing activities provided by the CoAA gene cluster will balance Tau isoforms and are important for neurodegenerative disease. Supporting the notion that alternative splicing contributes to both disease and cancer, we have previously found that the human CoAA gene cluster is amplified with recurrent loss of the 5′ regulatory enhancer in many common solid tumors (28). The enhancer loss associated with gene amplification prevents a complete alternative splicing switch from CoAA to CoAM at the stage of embryoid body cavitation during normal stem cell differentiation (27). Our current study suggests the presence of complex alternative splicing and trans-splicing patterns within the CoAA gene cluster encoding transcriptional coregulators. When differentially expressed, these coregulator variants will facilitate lineage-specific gene expression during cell differentiation and development. A defective control of alternative splicing in the CoAA gene cluster potentially leads to disease and cancer.

Acknowledgments—We thank Lori Redmond, Yu-chih Lin, and Nahid F. Mivechi for valuable help. We thank Doris B. Cawley for immunohistochemistry analysis.

REFERENCES

1. Berget, S. M., Moore, C., and Sharp, P. A. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 3171–3175
2. Chow, L. T., Gelinas, R. E., Broker, T. R., and Roberts, R. J. (1977) Cell 12, 1–8
3. Modrek, B., and Lee, C. (2002) Nat. Genet. 30, 13–19
4. Sanford, J. R., and Caceres, J. F. (2004) J. Cell Sci. 117, 6261–6263
5. Pan, Q., Shai, O., Lee, L. J., Frey, B. J., and Blencowe, B. J. (2008) Nat. Genet. 40, 1413–1415
6. Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P., and Burge, C. B. (2008) Nature 456, 470–476
7. Stetefeld, J., and Ruegg, M. A. (2005) Trends Biochem. Sci. 30, 515–521
8. Stamm, S., Ben-Ari, S., Rafalska, I., Tang, Y., Zhang, Z., Toiber, D., Thanaraj, T. A., and Soreq, H. (2005) Gene 344, 1–20
9. Smith, C. W., Patton, J. G., and Nadal-Ginard, B. (1989) Annu. Rev. Genet. 23, 527–577
10. Kalnina, Z., Zayakin, P., Silina, K., and Liné, A. (2005) Genes Chromosomes Dev. 18045
trans-Splicing Variants of CoAA and RBM4

11. Venables, J. P. (2006) BioEssays 28, 378–386
12. Xing, Y., and Lee, C. (2006) Nat. Rev. Genet. 7, 499–509
13. Mayer, M. G., and Floeter-Winter, L. M. (2005) Mem. Inst. Oswaldo Cruz 100, 501–513
14. Akiva, P., Toporik, A., Edelweiter, S., Peretz, Y., Diber, A., Shemesh, R., Novik, A., and Sorek, R. (2006) Genome Res. 16, 30–36
15. Eddy, S. R. (2001) Nat. Rev. Genet. 2, 919–929
16. Mattick, J. S., and Makunin, I. V. (2006) Hum. Mol. Genet. 15, R17–29
17. Brose, M. S., Volpe, P., Paul, K., Stopfer, J. E., Colligon, T. A., Calzone, K. A., and Weber, B. L. (2004) Genet. Test 8, 133–138
18. Neklason, D. W., Solomon, C. H., Dalton, A. L., Kuwada, S. K., and Burt, R. W. (2004) Fam. Cancer 3, 35–40
19. Ozcelik, H., Nedelcu, R., Chan, V. W., Shi, X. H., Murphy, J., Rosen, B., and Andrulis, I. L. (1999) Hum. Mutat. 14, 540–541
20. Adamia, S., Reiman, T., Crainie, M., Mant, M. J., Belch, A. R., and Pilarski, L. M. (2005) Blood 105, 4836–4844
21. Yeo, G. W., Xu, X., Liang, T. Y., Muotri, A. R., Carson, C. T., Coufal, N. G., and Gage, F. H. (2007) PLoS Comput. Biol. 3, 1951–1967
22. Boutz, P. L., Stoilov, P., Li, Q., Lin, C. H., Chawla, G., Ostrow, K., Shiue, L., Ares, M., Jr., and Black, D. L. (2007) Genes Dev. 21, 1636–1652
23. Auboeuf, D., Dowhan, D. H., Li, X., Larkin, K., Ko, L., Berget, S. M., and O’Malley, B. W. (2004) Mol. Cell. Biol. 24, 442–453
24. Kornblihtt, A. R. (2005) Curr. Opin. Cell Biol. 17, 262–268
25. Iwasaki, T., Chin, W. W., and Ko, L. (2001) J. Biol. Chem. 276, 33375–33383
26. Yang, Z., Sui, Y., Xiong, S., Liour, S. S., Phillips, A. C., and Ko, L. (2007) Nucleic Acids Res. 35, 1919–1932
27. Liu, F., and Gong, C. X. (2008) Mol. Neurodegener. 3, 8
28. D’Souza, I., and Schellenberg, G. D. (2005) Biochim. Biophys. Acta 1739, 104–115
29. Goedert, M. (2005) Mov. Disord. 20, Suppl. 12, S45–52
30. Pfuhl, T., Mamiani, A., D¨urr, M., Welter, S., Steiber, J., Ankara, J., Liss, M., Dobner, T., Schmitt, A., Falkai, P., Kremmer, E., Jung, V., Barth, S., Gr¨asser, F. A. (2008) Neurosci. Lett. 444, 11–15
31. Krause, M., and Hirsh, D. (1987) Cell 49, 753–761
32. Tessier, L. H., Keller, M., Chan, R. L., Fournier, R., Weil, J. H., and Imbault, P. (1991) EMBO J. 10, 2621–2625
33. Rajkovic, A., Davis, R. E., Simonsen, J. N., and Rottman, F. M. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 8879–8883
34. Dorn, R., Reuter, G., and Loewendorf, A. (2001) Proc. Natl. Acad. Sci. U.S.A. 98, 9724–9729
35. Eul, J., Graessmann, M., and Graessmann, A. (1995) EMBO J. 14, 3226–3235