Nuclear matrix-associated protein SMAR1 regulates alternative splicing via HDAC6-mediated deacetylation of Sam68

Kiran Kumar Nakka1,a, Sidhi Chaudharyd, Shruti Joshib, Jyotsna Bhatb, Kulwant Singhb, Subhrangsu Chatterjeec, Renu Malhotrab, Abhijit Dea, Manas Kumar Santrae, F. Jeffrey Dilworthb,f, and Samit Chattopadhyaya1,a

aChromatin and Disease Biology Laboratory, National Centre for Cell Science, Pune 411007, India; bSprott Centre for Stem Cell Research, Ottawa Hospital Research Institute, Ottawa, ON, K1H 8L6, Canada; cDepartment of Biophysics, Bose Institute, Kolkata 700094, India; dAdvanced Center for Treatment, Education and Research, Tata Memorial Centre, Navi Mumbai 410210, India; eCancer and Epigenetics Lab, National Centre for Cell Science, Pune 411007, India; and fDepartment of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON, K1H 8L6, Canada

Edited by Mark Groudine, Fred Hutchinson Cancer Research Center, Seattle, WA, and approved May 15, 2015 (received for review September 29, 2014)

Pre-mRNA splicing is a complex regulatory nexus modulated by various trans-factors and their posttranslational modifications to create a dynamic transcriptome through alternative splicing. Signal-induced phosphorylation and dephosphorylation of trans-factors are known to regulate alternative splicing. However, the role of other posttranslational modifications, such as deacetylation/acyetylation, methylation, and ubiquitination, that could modulate alternative splicing in either a signal-dependent or -independent manner remain enigmatic. Here, we demonstrate that Scaffold/matrix-associated region-binding protein 1 (SMAR1) negatively regulates alternative splicing through histone deacetylase 6 (HDAC6)-mediated deacetylation of RNA-binding protein Sam68 (Src-associated substrate during mitosis of 68 kDa). SMAR1 is enriched in nuclear splicing speckles and associates with the snRNAs that are involved in splice site recognition. ERK–MAPK pathway that regulates alternative splicing facilitates ERK-1/2–mediated phosphorylation of SMAR1 at threonines 345 and 360 and localizes SMAR1 to the cytoplasm, preventing its interaction with Sam68. We showed that endogenously, SMAR1 through HDAC6 maintains Sam68 in a deacetylated state. However, knockdown or ERK-mediated phosphorylation of SMAR1 releases the inhibitory SMAR1–HDAC6–Sam68 complex, facilitating Sam68 acetylation and alternative splicing. Furthermore, loss of heterozygosity at the Chr.16q24.3 locus in breast cancer cells, wherein the human homolog of SMAR1 (BANP) has been mapped, enhances Sam68 acetylation and CD44 variant exon inclusion. In addition, tail-vein injections in mice with human breast cancer MCF-7 cells depleted for SMAR1 showed increased CD44 variant exon inclusion and concomitant metastatic propensity, confirming the functional role of SMAR1 in regulation of alternative splicing. Thus, our results reveal the complex molecular mechanism underlying SMAR1-mediated signal-dependent and -independent regulation of alternative splicing via Sam68 deacetylation.

SMAR1 | Sam68 | HDAC6 | alternative splicing | deacetylation

Newly synthesized pre-mRNAs undergo multiple posttranscriptional gene-regulatory events, such as capping, splicing, cleavage, and polyadenylation. Of these, splicing is most stringently regulated, because it is a prerequisite for spatiotemporal generation of splice variants observed in 95% of human genes (1). Accuracy of alternative splicing (AS) is modulated by cis-regulatory elements that include splice enhancers and silencers and/or trans-acting factors, viz. the serine–arginine (SR)-rich family proteins, heteronuclear ribonucleoproteins (hnRNPs) and various chromatin modifiers (2, 3). The majority of these trans-factors are distributed along the nuclear matrix (NM), which is a fibro-granular structural framework, organized as chromatin and ribonucleoprotein (RNP) domains. Chromatin domains are involved in replication and transcription, whereas RNP domains are actively associated with cotranscriptional and post-transcriptional gene-regulatory events (4, 5). Thus, NM is considered to be the site for pre-mRNA processing; however, the precise role of NM-associated proteins in the regulation of AS remain elusive.

Scaffold/matrix-associated region-binding protein 1 (SMAR1) is one such NM-associated protein identified in mouse double-positive thymocytes (6). The human homolog of SMAR1, BANP, is mapped to chromosome 16q24.3 locus. Loss of heterozygosity (LOH) in this locus is implicated in multiple cancers (7, 8). SMAR1 delays tumor progression (9, 10) and inhibits the expression of multiple genes (11, 12). Clinical specimens of different histological grade breast cancer tissues indicated that SMAR1 expression is drastically down-regulated in the higher grades of infiltrating ductal carcinomas (13).

Cancer cells often take advantage of AS because the spliceosome machinery goes berserk in cancers, resulting in the expression of protein isoforms that promote growth and metabolism of tumor cells (14–16). Signal-transduction pathways—such as Ras, PI3-kinase, AKT, ERK, PRP4, and SRPK1—have been reported to be deregulated in cancerous cells (17). Posttranslational modifications (PTMs) induced by kinases and phosphatases on regulatory factors modulate tumor progression and metastasis, as exemplified by CD44 gene AS (18–21). CD44, a cell-surface transmembrane glycoprotein, expresses variant isoforms (CD44v) that increase

Significance

Multiple studies highlight the role of various proteins in regulation of alternative splicing; however, the regulatory role of distinct posttranslational modifications during alternative splicing that contribute to tumorigenesis is enigmatic. Here we report a previously unidentified noncanonical mechanism of regulation of alternative splicing modulated by deacetylation of RNA-binding protein Sam68 (Src-associated substrate during mitosis of 68 kDa) via Scaffold/matrix-associated region-binding protein 1 (SMAR1)–histone deacetylase 6 (HDAC6) complex. SMAR1 in complex with HDAC6 maintains Sam68 in a deacetylated state. We observed that ERK-1/2–dependent phosphorylation of SMAR1, knockdown of SMAR1, or loss of heterozygosity facilitates CD44 variant exon inclusion via Sam68 acetylation and thus confers invasive and metastatic potential in breast tumor cells. Our findings provide key insights into regulation of alternative splicing and the potential for therapeutic intervention during tumor metastasis.


The authors declare no conflict of interest.

This article is a PNAS Direct Submission. Freely available online through the PNAS open access option.

1To whom correspondence should be addressed. Email: samit@nccs.res.in.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1418603112/-/DCSupplemental.
metastatic propensity in multiple tumor types (22). Generation of CD44v is modulated by AS events, which require activation of the ERK–MAPK pathway (23). However, a comprehensive role for trans-factors that negatively influences CD44 AS in a signal-dependent and -independent manner remains to be established.

Our present study revealed SMAR1 as a negative regulator of CD44 gene AS in a signal-dependent and -independent manner by modulating Sam68 (Src-associated substrate during mitosis of 68 kDa) deacetylation. Here, we document that SMAR1, in cooperation with histone deacetylase 6 (HDAC6), interacts with Sam68 and maintains it in a deacetylated state, concomitantly inhibiting the inclusion of CD44 alternate exons. Furthermore, we show that ERK activation, knockdown of SMAR1, LOH at chromosome 16q24.3 locus, and/or inactivation of HDAC6 enhances Sam68 acetylation and facilitates CD44 variant exon inclusion. Overall, our results demonstrate the role of NM-associated protein SMAR1 in the regulation of AS.

Results

SMAR1 Is Enriched in Splicing Speckles and Regulates AS. SC35-enriched nuclear speckles are considered as a hallmark of proteins involved in pre-mRNA splicing (24). Presence of a stretch of serine and proline residues, in addition to interspersed arginine residues in the protein-interaction domain of SMAR1 (SI Appendix, Fig. S1A), inspired us to examine the distribution of SMAR1 in the RNP matrix. Initially, we found that the highly abundant SMAR1 protein had a homogenous distribution within the nucleus, making it difficult to resolve its distribution in the RNP matrix and/or its association with SC35-splicing speckles using immunofluorescence (IF) of intact cells. However, in situ immunostaining of NM (25) allowed us to reveal that SMAR1 is enriched in the RNP-rich splicing speckles and colocalizes with SC35 (Fig. 1A, –RNase A). Treatment of NM with RNase A abolished the occurrence of these nuclear speckles (Fig. 1A, +RNase A), suggesting that SMAR1-enriched speckles are RNPs in nature. Further, RNA immunoprecipitation (RIP) assays showed an association of SMAR1 with splicing component snRNAs, U1, U2, U4, U5, and U6 snRNAs that are involved in splice site recognition (Fig. 1B). Interestingly, affinity of SMAR1 for U5 snRNA is higher than it is for any other snRNA. The specific association of SMAR1 with snRNAs, but not with GAPDH mRNA, prompted us to investigate its effect on the regulation of AS. In vivo splicing assays with minigenes harboring the cis-regulatory elements controlling the AS of either FAS (26) or CD44 (23) revealed that SMAR1 knockdown augments the inclusion of alternative exons, whereas overexpression reduced the exon inclusion (Fig. 1 C and D and SI Appendix, Fig. S1 B and C). However, to eliminate the possibility that overexpression or knockdown of SMAR1 alters the stability of mRNA splice variants per se, we performed mRNA stability assays using Actinomycin D. Inhibition of transcription by using Actinomycin D in HEK293T cells that were previously cotransfected with either SMAR1 overexpression or knockdown plasmids along with FAS and CD44 minigenes resulted in no change in the observed ratio of spliced mRNAs at any given time point, indicating that SMAR1 expression does not contribute toward the differential stability of mRNA splice variants (Fig. 1 E and F). Together, these data indicate that SMAR1 specifically regulates AS without affecting the stability of splice variants.

SMAR1 Knockdown or LOH at Chr.16q24.3 Enhances CD44 Variant Exon Inclusion and Promotes Lung Metastasis. Considering that the levels of SMAR1 are drastically down-regulated in the higher grades of breast tumors (13), and the inclusion of CD44 variant exons confers invasive and metastatic propensity to tumor cells (27, 28), we set out to decipher the molecular mechanism underlying SMAR1-mediated regulation of CD44 AS. Semiquantitative exon-specific RT-PCR analysis of CD44 splice variants using different primer sets (20) (Fig. 2 A and SI Appendix, Table S1) in HeLa cells that are stably expressing a SMAR1-targeting shRNA (Fig. 2B) revealed an enhanced incorporation of variant exons v2–v6 (proximal variant exons; Fig. 2C). Interestingly, we observed no significant change in the incorporation of variable exons v7–v10 (distant variant exons; SI Appendix, Fig. S1D, Left) and standard CD44 (CD44s). In corollary, overexpression of SMAR1 (Fig. 1D) inhibited the inclusion of proximal variant exons (Fig. 1E), whereas distal variant exons and CD44s remain unaltered.
indicating that SMAR1 negatively regulates the inclusion of CD44 variable exons v2–v6.

Human breast cancer cells such as MDA-MB-468 and T47D harbor LOH, whereas MCF-7 cells lack LOH (7, 29). Western blot analysis showed that MDA-MB-468 and T47D cells express minimal amounts of SMAR1, whereas MCF-7 cells express significant amounts of SMAR1, corroborating with the LOH status (Fig. 2F). Because SMAR1 expression inversely correlates with CD44 variant exon inclusion, we investigated CD44 AS in context of LOH. In MDA-MB-468 and T47D cells, we observed a categorical increase in the inclusion of variant exons v2–v10, whereas CD44 variant exon inclusion is negligible in SMAR1 expressing MCF-7 cells (Fig. 2G and SI Appendix, Fig. S1E, Left). Nevertheless, ectopic expression of SMAR1 in LOH-harboring MDA-MB-468 and T47D cell lines by using adeno-SMAR1 (Ad-SM) (Fig. 2H) consistently decreased the inclusion of proximal exons (Fig. 2I), whereas no significant change was observed in the levels of distal variant exons and CD44s (SI Appendix, Fig. S1E, Right). Together, these data indicate that LOH at 16q24.3 locus, wherein SMAR1 has been mapped, contributes to CD44 AS.

Considering that the inclusion of CD44 variable exons confers the invasive and metastatic proclivity to tumor cells, we investigated the metastatic potential of MCF-7SMAR1−/− cells. Tail-vein injections in NOD/SCID mice with MCF-7 cells that were previously stably transduced with nonsilencing (NS; n = 6) and SMAR1–shRNA lentivirus (sh3; n = 6) lentivirus. Lung metastasis assays were performed, and lung nodules are indicated by arrows (J, Upper Left) and were quantified (n = 3) as shown in the graph (J, Right). Histopathological analysis [hematoxylin/eosin (HE) staining] of the lung sections (J, Lower Left). Student’s t tests were used to calculate the P values. (K) Immunohistochemistry (IHC) analysis of the lung sections (n = 3) obtained from metastatic breast tumor nodule sites with the indicated proteins (α-SMAR1 and various α-CD44v antibodies).
(NOD/SCID) mice indicated that MCF-7\textsuperscript{SMAR1−/−} cells showed a profound increase in invasive and metastatic potential, especially to the lungs, as evidenced by histopathological analysis (Fig. 2J). Immunostaining of the histological sections of metastatic lung nodules with α-SMAR1 and -CD44 variable exon-specific antibodies indicated enhanced expression of CD44v in the metastatic nodules, wherein the amount of SMAR1 is minimal (Fig. 2K). All together, these results indicate that knockdown or absence of SMAR1 enhances CD44 alternate exon inclusion, and the metastatic propensity served as a hallmark of CD44v inclusion.

**Activation of ERK–MAPK Pathway Translocates SMAR1 to the Cytoplasm.**

ERK–MAPK signaling is considered indispensable and a prerequisite for CD44 gene AS (18, 23). Considering the similar effects exerted on CD44 gene AS upon low levels of SMAR1 and ERK–MAPK signal activation, we predicted an intricate interplay between SMAR1 expression and ERK–MAPK pathway. Towards this prediction, activation of ERK–MAPK pathway using phorbol 12-myristate 13-acetate (PMA) in HeLa cells showed no change in SMAR1 levels (SI Appendix, Fig. S2). However, IF analysis revealed that SMAR1, a primarily nuclear protein, translocated to the cytoplasm upon ERK activation (Fig. 3A). This observation was reaffirmed by Western blot analysis of SMAR1 in PMA-stimulated nuclear-cytoplasmic fractions (Fig. 3B). To further determine that SMAR1 localization is ERK-dependent, we used ERK-1/2 inhibitor U0126 and observed an abrogation in the translocation of SMAR1 to the cytoplasm (Fig. 3A and C). Furthermore, treatment with a nuclear export inhibitor, Leptomycin B, also inhibited the cytoplasmic translocation of SMAR1 (Fig. 3A and D), reinforcing that activated ERK-1/2 kinase induces nuclear export of SMAR1.
Collectively, these results indicate that ERK–MAPK signal mimics SMAR1 knockdown by facilitating its cytoplasmic localization, thus promoting CD44 gene AS.

**SMAR1 Is Accumulated Along the CD44 Gene Body and Binds to CD44 Pre-mRNA.** An intricate association of chromatin remodelers with splicing machinery and the chaperoning of pre-mRNA by chromatin during AS incubed us to analyze the distribution of SMAR1 along the CD44 gene locus (Fig. 3E). Chromatin immunoprecipitation (ChIP) assays indicated an endogenous distribution of SMAR1 along the intronic and exonic regions of CD44; however, activation of the ERK–MAPK pathway reduced the same (Fig. 3F). Furthermore, occupancy of RNA polymerase (Pol) II (19) inversely correlates with that of SMAR1 and also served as positive control (SI Appendix, Fig. S3 A and B). Considering that some of the chromatin proteins, such as heterochromatin protein 1γ (HP1γ) and RNA Pol II, bridge the nascent pre-mRNA to the chromatin (21), we further investigated the association of SMAR1 with CD44 pre-mRNA. RIP assays revealed an intricate association of SMAR1 with CD44 pre-mRNA in untreated cells (Fig. 3G); however, the association was reduced upon ERK-1/2 activation (Fig. 3G, red line). The association of SMAR1 with pre-mRNA containing variable exons v3 and v5 was significantly high, indicating that SMAR1 associates with CD44 transcript and likely sequesters the variable exons from being incorporated into the mRNA.

**Activated ERK-1/2 Kinase Phosphorylates SMAR1 at Threonines 345 and 360.** ERK–MAPKs are classical protein kinases that phosphorylate numerous substrates with proline-directed serine/threonine residues (18, 30). SMAR1 in its protein-association domain harbors two putative threonine followed by proline (TP) residues positioned at 345 and 360 (Fig. 4A). Phospho-TP (pTP) assays using α-pTP antibody revealed time-dependent phosphorylation of SMAR1 upon ERK-1/2 activation (Fig. 4B). SMAR1 phosphorylation at TP residues was observed within 15 min of ERK activation, and the levels in the nucleus decreased by 60 min (Fig. 4B, lanes 2 and 5), indicating nuclear export of phosphorylated protein. A concomitant increase of pTP–SMAR1 in the cytoplasmic fractions was observed within 15 min until 120 min of ERK activation. However, treatment with U0126 inhibited the phosphorylation of SMAR1 as well as its cytoplasmic localization (SI Appendix, Fig. S4A). Furthermore, in vitro kinase assay, by incubating recombinant GST–SMAR1 with activated recombinant ERK2 in a concentration-dependent manner, indicated phosphorylation of SMAR1 (Fig. 4C). However, treatment with calf intestinal phosphatase (CIP) abated SMAR1 phosphorylation, affirming SMAR1 as a substrate of activated ERK kinase (Fig. 4D). To further ascertain these observations, we generated threonine-to-alanine mutants (T345A/T360A) in p3X-Flag-SMAR1. IF (Fig. 4E) and pTP assays (Fig. 4F), along with subcellular fractionation studies (SI Appendix, Fig. S4B), indicated that T345A/T360A mutants are phosphorylation-deficient and failed to exhibit cytoplasmic localization, rendering insensitivity to ERK activation. To further strengthen these observations, we generated in-house antibodies against T345/T360 phospho-SMAR1 (Pep3) and T345A/T360A mutant-SMAR1 (Pep6) (SI Appendix, Fig. S4C). Analysis of PMA-stimulated nuclear and cytoplasmic fractions using these antibodies reaffirmed phosphorylation-mediated differential localization of SMAR1 (SI Appendix, Fig. S4D). Despite Pep6-antibody-detected endogenous unphosphorylated SMAR1, which is primarily localized in the nucleus, it failed to recognize the cytoplasmic as well as nuclear threonine-phosphorylated SMAR1. These results confirm the specificity of the antibody as well as phosphorylation of SMAR1 at T345 and T360.

**Fig. 4.** ERK-1/2 kinase phosphorylates SMAR1 at T345 and T360. (A) Schematic representation of putative ERK-mediated SMAR1 phosphorylation sites at T345 and T360 (highlighted in red). (B) In vivo pTP analysis of α-SMAR1 eluates by immunoblotting with α-pTP antibody in both the nuclear and cytoplasmic fractions of HeLa cells activated with PMA for the indicated time points. IP, immunoprecipitation. (C) In vitro phosphorylation of recombinant GST–SMAR1 protein upon incubation with 20 or 80 units (U) of recombinant ERK2 in the presence of ATP. GST served as a control substrate. (D) In vitro dephosphorylation of GST–SMAR1 by calf intestinal phosphatase (CIP, 100 U). (E) IF analysis of HeLa cells transfected with FSM wild-type (FSM) or T345A/T360A mutant (FSM [Mut]) plasmids and stained with Flag (red, indicated by arrow) and DNA (DAPI). (F) In vivo pTP assays of α-Flag–immunoprecipitated eluates from the nuclear and cytoplasmic fractions of HeLa cells transfected with the indicated plasmids and treated with PMA (30 min). (G) RT-PCR monitoring of CD44 alternate exon inclusion in HeLa cells transfected with the indicated plasmids and left untreated or treated with PMA. (H) Western blot analysis of whole-cell lysate and nuclear and cytoplasmic fractions from HeLa cells transfected with the indicated plasmids.
To further demonstrate the biological significance of T345/T360 phosphorylation of SMAR1 and CD44 AS, we examined the CD44 variant exon incorporation upon expression of wild-type Flag-SMAR1 (FSM) and T345A/T360A mutant Flag-SMAR1 (FSM-Mut). As depicted in Fig. 4G, unactivated cells expressing FSM, as well as FSM-Mut, showed decreased CD44v inclusion. Interestingly, activation of the ERK pathway enhanced variant exon inclusion in cells expressing either vector alone or FSM, but not in cells expressing FSM-Mut. This result can be attributed to the presence of FSM-Mut in the nucleus (Fig. 4H), which impedes the ability of ERK-1/2 to induce a switch in CD44 variable exon inclusion. Together, these results indicate that phosphorylation of SMAR1 at T345/T360 is critical and necessary for ERK–MAPK signal-induced inclusion of CD44 variant exons.

**SMAR1 Knockdown Promotes AS Independent of the ERK–MAPK Pathway.** Given that SMAR1 knockdown facilitates CD44 gene AS in a manner similar to that of activation of the ERK–MAPK pathway (23), we investigated a possible leaky activation of ERK-1/2 upon SMAR1 knockdown. Toward this hypothesis, we checked for the expression of phosphorylated ERK-1/2 (activation of ERK) in the nuclear extracts from SMAR1 knockdown HeLa cells. Interestingly, unlike the PMA-stimulated nuclear extracts that showed the presence of phospho-ERK-1/2 (Fig. 5A, lane 2), SMAR1-knocked-down nuclear extracts did not show any of its expression (Fig. 5A, lane 4), indicating ineffective activation of ERK-1/2. However, we observed a consistent increase in CD44 variant exon inclusion upon SMAR1 depletion (Fig. 5B, lane 4), indicating that SMAR1 knockdown-mediated regulation of the CD44 gene AS is ERK–MAPK signal-independent. This finding incited us to investigate
the molecular mechanism underlying SMAR1-mediated signal-independent regulation of AS.

**SMAR1 Interacts with Sam68 in a Signal-Dependent Manner.** Sam68, a STAR-family protein, associates with splice-regulatory elements of pre-mRNAs and plays a strategic role in gene AS (18). Of significance, association with CD44 pre-mRNA prompted us to investigate its association with Sam68. Immunoprecipitation (IP) assays in HeLa cell nuclear extracts revealed an endogenous association of SMAR1 with Sam68 (SI Appendix, Fig. S5A). However, the association of SMAR1 with Sam68 is prevented in ERK-activated nuclear extracts due to the differential localization of SMAR1 to the cytoplasm (Fig. 5C, lanes 2 and 6). Furthermore, IP analysis of nuclear extracts treated with RNase A abolished the SMAR1–Sam68 interaction. (SI Appendix, Fig. S5B). Together, these data suggest that endogenously, SMAR1 associates with Sam68 in an RNA-dependent manner, and this association is perturbed due to differential localization of SMAR1 upon activation of ERK-1/2. This finding led us to hypothesize that SMAR1 in complex with Sam68 might form an inhibitory complex that perturbs AS, which inspired us to hypothesize that SMAR1 in complex with Sam68 might form an inhibitory complex that perturbs AS, which inspired us to hypothesize that SMAR1 might modulate the acetylation/deacetylation dynamics of Sam68. IP assays using α-acetyl-lysine (α-AcK) antibody revealed enhanced acetylation of Sam68 upon SMAR1 knockdown (Fig. 5F, lanes 2 and 3), despite total Sam68 levels remaining unaltered, emphasizing that endogenously Sam68 is maintained in a deacetylated state and knockdown of SMAR1 facilitates Sam68 acetylation. Because ERK activation prevents SMAR1–Sam68 association, we investigated Sam68 acetylation upon ERK activation. Consistently, we observed an accumulation of acetylated Sam68 upon ERK activation (Fig. 5F, lanes 3–6). Of significance, acetylation kinetics of Sam68 indicated that phosphorylation of SMAR1 (Fig. 4B, 15 min) precedes Sam68 acetylation upon ERK activation (Fig. 5F, 30 min). Strikingly, inhibition of ERK-1/2 with U0126 prevented Sam68 acetylation (SI Appendix, Fig. S5F). Furthermore, Sam68 acetylation in SMAR1 knockdown cells treated with U0126 convincingly demonstrated that Sam68 acetylation depends on the presence or absence of SMAR1 in the nucleus (SI Appendix, Fig. S5G). Together, these data reveal that knockdown of SMAR1 and/or activation of ERK-1/2 are equally effective in facilitating Sam68 acetylation.

To further establish the significance of Sam68 acetylation during AS, we generated various deletion constructs of Sam68 (Fig. 5G) and verified their acetylation status as well as the efficiency to modulate AS. Results showed that myc-Sam68 full-length (1–443), myc-Sam68 (90–443), and myc-Sam68 (90–257) are acetylated upon ERK activation (Fig. 5H, lanes 2, 4, and 8), whereas myc-Sam68 (257–443), lacking lysine-rich KH domain, failed to show any acetylation (Fig. 5H, lane 6). Acetylation of the minimal domain, myc-Sam68 (90–257), harboring 21 lysine residues, but not myc-Sam68 (257–443), emphasized that lysine residues are the target sites of Sam68 acetylation. CD44 splice reporter minigene splicing assays in MCF-7/Samo68 cells (SI Appendix, Fig. S5H) indicated that the lysine-rich KH domain is indispensable for CD44 AS (SI Appendix, Fig. S5I). In this investigation, we assayed FAS minigene AS (SI Appendix, Fig. S5J) and one of the target genes of Sam68, Bel-X gene AS (33) (SI Appendix, Fig. S5K). Our results from these studies indicated that acetylation of Sam68 through the mechanism of ERK activation facilitates AS. Together, these results distinctly highlight that deacetylation/acetylation switch on Sam68 is critical for a regulatory event during AS.

**HDAC6 in Complex with SMAR1 Deacetylates Sam68 and Regulates AS.** Next, we investigated the specific deacetylate of Sam68 that coordinates with SMAR1 and regulates AS. Our results revealed that, despite SMAR1 association with both HDAC1 (21) and Sam68, Sam68 failed to interact with HDAC1 (SI Appendix, Fig. S6 A–C). However, a study by Wang et al. on genome-wide occupancy of HDACs revealed the enrichment of HDAC2 and HDAC6 on to the coding regions of multiple genes (34). An association of SMAR1 with HDAC6 (35), but not with HDAC2 (SI Appendix, Fig. S6 D and E), and an association of HDAC6 with RNA Pol II (33) and with U5 snRNA (SI Appendix, Fig. S6F) indicated that SMAR1 might cooperate with HDAC6 to deacetylate Sam68. Consistent with our hypothesis, IP assays clearly showed that SMAR1, as well as Sam68, associate with HDAC6 (Fig. 6A). Concurrently, sequential IP experiments revealed the coexistence of SMAR1–HDAC6–Sam68 as a trimeric complex (Fig. 6B). Furthermore, the existence of this trimeric complex was confirmed by size-exclusion chromatography using HeLa cell nuclear extracts (SI Appendix, Fig. S6G, Upper, fractions 16–26). However, the trimeric complex was observed to be RNA-sensitive because RNase A-treated nuclear extracts showed the dissociation of high-molecular-weight HDAC6–SMAR1 and Sam68 complexes (SI Appendix, Fig. S6G, Lower, fractions 16–26). Thus, in an in silico analysis of HDAC6–SMAR1–Sam68 complex in the presence of RNA component is compatible with the positioning of RNA at the interacting site of SMAR1–Sam68, and both the electrostatic and solution energies contribute largely for the stabilization of protein–trimeric and RNA–protein–trimeric complexes (Fig. 6C and SI Appendix, Table S3). Furthermore, positioning of the RNA enhances the total energy stabilization of the protein–trimeric complex (SI Appendix, Table S4). Based on the above observations, it can be inferred that SMAR1, through HDAC6, maintains Sam68 in a deacetylated state.

To prove that HDAC6 plays a decisive role in the modulation of Sam68 acetylation and in turn CD44 gene AS, we investigated the consequences upon HDAC6 knockdown. Semiquantitative RT-PCR analysis of CD44 splice variants in HeLa cells that were stably knocked down with HDAC6 shRNA lentivirus revealed an increase in the incorporation of CD44 splice variants v3, v5, v6, and v7 (Fig. 6D). Conversely, we observed enhanced acetylation of Sam68 upon HDAC6 knockdown, indicating that HDAC6 maintains Sam68 in a deacetylated state (Fig. 6E). Results of Sam68 acetylation (SI Appendix, Fig. S6H) and CD44 gene AS (SI Appendix, Fig. S6I) upon knockdown of HDAC6 showed no change in Sam68 acetylation or AS, indicating that HDAC6 is the specific deacetylase that coordinates with SMAR1 for Sam68 deacetylation. Nevertheless, to further extrapolate and establish the biological significance of SMAR1 in regulating HDAC6-mediated deacetylation of Sam68 (and subsequent AS of the CD44), we investigated Sam68 acetylation in MDA-MB-468, T47D, and MCF-7 breast cancer cells. Of significance, we observed acetylation of Sam68 in LOH-harboring MDA-MB-468 and T47D cells (Fig. 6F, lanes 2 and 3), which in the earlier results were shown to exhibit enhanced incorporation of CD44 variable exons (Figs. 2F and 2G). Sam68 remained deacetylated in MCF-7 cells that do not harbor LOH. Strikingly, MDA-MB-468 and T47D cells expressed higher HDAC6 levels than MCF-7 cells (Fig. 6F). This finding highlights the role of SMAR1 in the maintenance of Sam68 in a deacetylated state. IP assays in MCF-7 cells, which express SMAR1, maintain HDAC6–SMAR1–Sam68 complex; however, the interaction was found to be perturbed in MDA-MB-468 and T47D (Fig. 6G, Left). Notably, the interaction of HDAC6 with Sam68 was restored upon...
SMAR1 expression in LOH-harboring MDA-MB-468 and T47D cells (Fig. 6G, Middle and Bottom Right). Contrastingly, MCF-7 cells that were knocked down for SMAR1 failed to show the interaction between HDAC6 and Sam68 (Fig. 6G, Upper Right). Together, these data convincingly establish that SMAR1 is a critical mediator of the HDAC6–Sam68 interaction and that SMAR1 coordinates HDAC6-mediated deacetylation of Sam68.

Inactivation of HDAC6 Facilitates Sam68 Acetylation and AS. To further ascertain our observations of HDAC6-mediated deacetylation of Sam68 and CD44 gene AS, we used an HDAC6-specific inhibitor, tubacin. We observed that HeLa cells with tubacin resulted in the enhanced incorporation of CD44 alternate exons v3–v7 (Fig. 6H). Furthermore, analysis of nuclear extracts from tubacin-treated HeLa cells revealed distinct acetylation of Sam68 (Fig. 6I). Considering that HDAC6 in complex with SMAR1 deacetylates Sam68 and to establish that LOH-mediated Sam68 acetylation is key to enhance CD44 variant exon inclusion (Figs. 6F and 2G), we treated MCF-7 cells with tubacin and examined Sam68 acetylation, as well as the CD44 variant exon inclusion. Consistent with our assumption, we observed a favorable switch in the incorporation of CD44 alternate exons v3–v7 (Fig. 6I), as well as Sam68 acetylation (Fig. 6K). However, the variables v5 and v6 showed a decrease in incorporation after 6 h of tubacin treatment. This result can be attributed to either the toxicity or altered stability of tubacin in these cells. Together, these results ascertain that HDAC6 deacetylates Sam68, and inactivation of HDAC6 favors Sam68 acetylation, which in turn facilitates AS of the CD44 mRNA.

Discussion
Our present study delineates a complex regulatory network underlying the signal-dependent and -independent regulation of cotranscriptional AS by SMAR1 (Fig. 7). Endogenously, SMAR1 in complex with HDAC6 associates with Sam68 and maintains it in a deacetylated state, consecutively inhibiting CD44 variant exon inclusion. Alternatively, activation of the ERK–MAPK pathway, which is considered to be indispensable for CD44 gene AS, disrupts the HDAC6–SMAR1–Sam68 complex, enhancing Sam68 acetylation, which results in cotranscriptional inclusion of CD44 gene variant exons. Similarly, knockdown of SMAR1, which also favors Sam68 acetylation independent of ERK–MAPK pathway signaling, enhances CD44 gene AS. Together, these results emphasize the significance of Sam68 acetylation in the context of regulated splicing.

SMAR1 is a scaffold/matrix attachment region-binding protein that regulates gene expression by binding to various promoter elements (11, 12). Here, we show that SMAR1 is enriched in the nuclear speckles, which are the sites for the association of actively transcribing genes, wherein chromatin modulators, splicing factors, snRNPs, and hnRNPs are aggregated (24). Association of SMAR1 with the spliceosome components U1, U2, U4, U5, and U6 snRNAs provided our initial insights into the role of SMAR1 in splice site selection. It is noted that certain other NM factors, such as SAFB1 and SAFB2, are also reported to modulate the alternative exon inclusion (36, 37).

Cancer cells often express aberrant splice variants due to altered expression of trans-factors. For instance, enhanced expression of the hnRNP A1/A2 and PTB represses the inclusion of exon 9 in the pyruvate kinase M gene, which contributes to the Warburg effect in cancer cells (16). Similarly, CD44 gene AS distinctly illustrates the role of variable exons during the metastasis of various tumors (22, 28). Our findings underscore the significance of SMAR1-mediated CD44 gene splicing in the context of breast cancer metastasis.
The ERK–MAPK pathway substantially enhances the inclusion of CD44 variant exons (23). Proteins such as Sam68, SRm160, Brm, and HP1γ that regulate CD44 pre-mRNA splicing are substrates of ERK kinase (18, 20, 21). Given that SMAR1 knockdown enhances CD44 variant exon inclusion, we observed phosphorylation-mediated translocation of SMAR1 from nucleus to the cytoplasm upon ERK activation, thus mimicking SMAR1 knockdown. As such, our data provide further insights into the mechanism through which the ERK–MAPK pathway induces CD44 variant exon inclusion. Mutation of FSM at T345A/T360A, which is phosphorylation-deficient, failed to translocate to the cytoplasm upon ERK activation, inhibiting CD44 alternate exon inclusion. ChIP and RIP assays that indicated association of SMAR1 with the chromatin and CD44 pre-mRNA highlighted an inverse correlation in the occupancy of Brm, HP1γ, and RNA Pol II upon ERK activation (19, 21). Because SMAR1 knockdown favors CD44 alternate exon inclusion, the possibility of increased occupancy of Brm and HP1γ under such a scenario, which in turn enhances the occupancy of serine 5 phospho-RNA Pol II, cannot be ruled out. This possibility might be one of the contributing factors that enhance CD44 gene AS upon SMAR1 knockdown, because RNA Pol II with decreased elongation rate facilitates the use of weaker splice sites (38).

Despite failed activation of ERK-1/2 upon SMAR1 knockdown, increased inclusion of CD44 alternate exons indicates the existence of a noncanonical pathway for CD44 gene AS. Because PTMs are known to play a key regulatory role during a majority of the cellular events, we investigated possible PTMs that can be further converted to acetylation-based events (38). Thus, our study is in agreement with this hypothesis and depicts that SMAR1-1/2-mediated phosphorylation of SMAR1 and Sam68 (18) is a prerequisite to disrupt their association, resulting in Sam68 acetylation during signal-dependent alternate exon inclusion. A study using various deletion and mutant constructs of Sam68 distinctly highlighted the role of Sam68 KH-domain residues and the requirement for Sam68 acetylation during AS of CD44. Because the pattern of acetylation alters the mechanistic impact, acetylation at a cluster of lysine residues results in the formation of charged patches (39), which in turn might enhance the affinity of Sam68 toward RNA. Together, Sam68 acetylation favors the inclusions of exons with weak splice sites and thus might be offering a fail-safe mechanism of CD44 gene AS during embryogenesis.

Deacetylation of Sam68 by SMAR1 in cooperation with HDAC6 highlights the fact that the SMAR1–HDAC1 pool, which is mostly localized at gene promoters, is distinct from SMAR1–HDAC6–Sam68 pool that localizes to the CD44 gene body and regulates AS. The basis for the specificity of HDAC6 association with Sam68 can be attributed to its differential occupancy. HDAC1 and HDAC3 are mostly localized to promoters, whereas HDAC2 and HDAC6 are enriched on the gene body as well as promoters (35). The association of HDAC6 with RNA Pol II (35) further contributes to the HDAC6-mediated deacetylation of Sam68. In this case, it is tempting to speculate that HDAC6 could also modulate the elongation rate of RNA Pol II via deacetylation of additional transcriptional elongation and splicing factors that are bound to the CTD tail, resulting in regulated splicing through an alternative mechanism.

This previously unidentified molecular mechanism of regulated AS via Sam68 acetylation adds an extra dimension in understanding cancer prognosis, especially in breast cancer. Our data, to our knowledge, for the first time show that LOH at 16q24.3 locus in breast cancer cells is associated with PTMs of Sam68 and splicing factors that are bound to the CTD tail, resulting in significantly low levels of CD44 variable exons. Despite the presence of relatively high levels of HDAC6 in MDA-MB-468 and T47D compared with MCF-7, our work suggests that Sam68 is maintained in an acetylated state in these cells due to their lack of SMAR1, which is indispensable for the colocalization of HDAC6 and Sam68. Despite decreased incorporation of proximal variant exons upon restoration of SMAR1 expression in LOH-harboring cells, the distal exons remain unaltered. This observation highlights that there are some other contributing factors encoded from the 16q24 locus, which might modulate the incorporation of CD44 distal variant exons. Alternatively, enhanced interaction of Sam68 with the U5 snRNA upon ERK activation (19) and endogenous association of SMAR1 with U5 snRNA indicates that in addition to modulation of Sam68 acetylation, cooperation between SMAR1 and HDAC6 might act as “roadblock” to sequester U5 snRNA from Sam68, thus preventing the splice site selection.

In summary, our study elucidated a previously unappreciated regulatory mechanism used to control alternate exon use, whereby the NM-associated protein SMAR1 and chromatin-modifying enzyme HDAC6 modulate the acetylation of splicing factors like Sam68 to modulate AS in response to extracellular cues (Fig. 7). To our knowledge, we show for the first time that...
HDAC6 plays a pivotal role in regulated splicing of CD44 via Sam68 deacetylation. Our findings reveal that knockdown or reduced levels of SMAR1 promotes cancer cell invasion and metastasis, whereas overexpression restricts cell invasion. This finding suggests that targeting Sam68 acetylation may provide alternative therapeutic strategies during breast cancer metastasis. Thus, we provide a previously unidentified mechanism of regulation of AS in the etiology of human breast cancer.

Materials and Methods

Additional materials and methods on plasmids, antibodies, chromatin immunoprecipitation, and RNA immunoprecipitation, as well as on the modeling and docking studies, are provided in SI Appendix, SI Materials and Methods.

Cells and Cell Culture. HEK 293T, HeLa, and MCF-7 cells were cultured in DMEM supplemented with 10% (vol/vol) FBS, 100 U/mL penicillin, and 100 μg/mL streptomycin (Invitrogen) at 37 °C and 5% CO2. MDA-MB-468 and T47D cells were grown in L-15 and RPMI medium under similar conditions.

Adenovirus and Lentivirus. SMAR1 adenovirus was used as described. For SMAR1, HDAC6, and Sam68 lentivirus generation, HEK 293T cells were cotransfected with pSPAX, pMD2.G, SMAR1–shRNA (Clone ID: V3LHS_17233; V3LHS_17235; V3LHS_17231; V3LHS_17237; V3LHS_34090; V3LHS_34089; VLSH_17238; VLSH_17229; VLSH_9927 for nonsilencing; lentiviral Open Biosystems), HDAC6–shRNA (Clone ID: TRCN0000004839 (sh1), TRCN0000004840 (sh2), or Sam68–shRNA (Clone ID: TRCN000000444–TRCN000000004; TRCN000000101). Indicated cell lines were transduced with a 1:1 mix of viral supernatant and growth medium. Stable cell lines were selected with 1.5 μg/mL puromycin (Sigma).

RNA Isolation and Semiquantitative RT-PCRs. RNA was isolated by using TRIzol (Invitrogen). Reverse-transcription was performed with 2 μg of RNA using M-MuLV reverse transcriptase (Invitrogen), Oligo dT (IDT), and RNase OUT (Invitrogen). PCRs were performed in 25 μL of reaction mixture by using Taq DNA polymerase (Chromus Biotech).

In Situ NM Staining. Nuclear matrices were prepared as mentioned in a well-established protocol from Nickerson et al. (25). IF analysis of the isolated matrices was performed by using a standard protocol after fixation with paraformaldehyde (12).

Whole-Cell Lysates and Nuclear and Cytoplasmic Fractionations. Indicated cells were lysed in whole-cell lysis buffer (50 mM Tris Cl, pH 7.4, 1% Nonidet P-40, 250 mM NaCl, 0.5 mM Na3VO4) supplemented with complete protease-inhibitor mixture (Roche). For nuclear and cytoplasmic fractionations, cell nuclei were eluted at 95 °C in IP buffer with SDS loading dye. The eluates were loaded on to the SDS/PAGE and immunoblotted with indicated antibodies.

ACKNOWLEDGMENTS. We thank Prof. Harald König for providing CD44 PETs-Luc plasmid and pcDNA3.1 myc-Sam68 plasmids, and Prof. Juan Valcárcel Juárez for the FAS minigene construct. We also thank Dr. Shekhar C. Mande, Director, National Centre for Cell Science. This work was supported by the Council of Scientific and Industrial Research, India; the Department of Biotechnology, India; and the Indian Council of Medical Research, Government of India (S. Chattopadhyay), and Canadian Institutes of Health Research (F.J.D.).

SI APPENDIX
SUPPLEMENTAL INFORMATION

for

“Nuclear matrix-associated protein SMAR1 regulates alternative splicing via HDAC6-mediated deacetylation of Sam68”

Kiran Kumar Nakka\textsuperscript{a,b}, Nidhi Chaudhary\textsuperscript{a}, Shruti Joshi\textsuperscript{a}, Jyotsna Bhat\textsuperscript{c}, Kulwant Singh\textsuperscript{b}, Subhrangsu Chatterjee\textsuperscript{c}, Renu Malhotra\textsuperscript{d}, Abhijit De\textsuperscript{d}, Manas Kumar Santra\textsuperscript{e}, F. Jeffrey Dilworth\textsuperscript{b,f} and Samit Chattopadhyay\textsuperscript{a,1}

\textsuperscript{a}Chromatin and Disease Biology Laboratory, National Centre for Cell Science, Pune 411007, India.
\textsuperscript{b}Sprott Center for Stem Cell Research, Ottawa Hospital Research Institute, Ottawa, Ontario K1H 8L6, Canada.
\textsuperscript{c}Department of Biophysics, Bose Institute, Kolkata 700054, India.
\textsuperscript{d}Advanced Center for Treatment, Education and Research, Tata Memorial Centre, Navi Mumbai 410210, India.
\textsuperscript{e}Cancer and Epigenetics Lab, National Centre for Cell Science, Pune 411007, India.
\textsuperscript{f}Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario, K1H 8L6, Canada.

\textsuperscript{1}Correspondence: Chromatin and Disease Biology Laboratory, National Centre for Cell Science, Pune University Campus, Ganeshkhind, Pune 411007, India.
Tel: 91-20-2570 8064; Fax: +91 20 2569 2259;
E-mail: samit@nccs.res.in
SI Materials and Methods.

Plasmids, antibodies and regents

p3X-Flag-SMAR1 plasmid was used for overexpression of SMAR1. For knockdown of SMAR1, pSIREN Retro Q-Z SMAR1-sh745 clone was used. The sequence details are mentioned in Sinha et al., 2010. Recombinant GST-SMAR1 protein was expressed in BL21 and used as described in Rampalli et al., 2005. For CD44 gene reporter splicing analysis, pET-v5-Luc was used (Kind gift from Harald König). PMA and U0126 were purchased from cell signaling technology and used at a concentration of 50 ng/ml and 25 µM, respectively. Leptomycin B was purchased from Sigma and used at a concentration of 10 nM, 10 min prior to PMA treatment. Tubacin was purchased from Sigma and used at a concentration of 5 µM. Antibodies anti-p44/42 (ERK-1/2) and anti-phospho p44/42 (p-ERK-1/2; Cell signaling technology) were used at a dilution of 1:1000 in 3% BSA in TBST. Anti-HDAC6 (Cell Signaling technology#2162) was used at a dilution of 1:1000 in TBST for western blotting, and Abcam (ab47181), for RNA immunoprecipitation. Anti-SMAR1 (Bethy laboratories) was used at a dilution of 1:4000 in 1% BSA in TBST. Anti-Lamin and anti-Paxillin (Santa Cruz) were used at a dilution of 1:2500 in TBST. Anti-actin (Sigma) was used at a dilution of 1:6000 in TBST. Anti-Flag antibody was purchased from Sigma and used at a dilution of 1:2000 in TBST. Secondary antibodies towards anti-mouse and anti-rabbit primary antibodies were purchased from Invitrogen and were used at a dilution of 1:5000 in 1% skimmmed milk in TBST. Antibodies for anti-RNA Pol II, anti-serine 5 phosphorylated RNA Pol II, anti-serine 2 phosphorylated RNA Pol II and anti-acetyl lysine were purchased from Millipore. All chemicals used were of molecular biology grade and purchased from either Sigma or USB. (BSA, bovine serum albumin; TBST, tris-buffered saline and Tween-20).
Generation of Sam68 truncations and mutants

cDNA3.1 myc-Sam68 (Kind gift from Harald König) was used as a template in PCR reaction to generate various truncations of Sam68 such as (90-443), (257-443) and (90-257). PCR amplified fragments were cloned in to pcDNA3.1 myc and the expression was verified by western blotting. The clones were sequenced using ABI sequencer and the expression was checked by western blotting with α-myc antibody (1:1000; Roche, 9e10 Clone)

Antibody cross-linking and immunoprecipitation (IP)

Majority of the IP experiments were done via covalently cross-linking the antibody with A/G sepharose beads (Pierce) to avoid non-specific binding and contamination of immunoglobulin in the immunoprecipitated protein eluates. Approximately 1µg of antibody was cross-linked with the 60 µl (50% slurry) of beads. Beads were washed thrice with ice-cold PBS and further incubated with desired antibody in IP buffer (PBS with 0.1% NP-40) containing protease inhibitor cocktail (Roche). For efficient and specific binding, the antibody-bead mixture was incubated overnight at 4ºC. Unbound antibody was removed by one wash of the antibody bead complex with IP buffer. Further, the antibody-bead complex was incubated with 10 mg/ml of Dimethyl pimelimidate (DMP) for 45 min at room temperature with rotation. Traces of DMP can be removed by quenching and washing with triethanolamine (TEA) of 50 mM and 200 mM respectively. The antibody-bead complex was further equilibrated with the IP buffer (0.1% NP-40 in PBS). These beads were either used immediately or stored at 4ºC for 2-3 days. Protein bound to the antibody-bead complex was eluted by using mild elution buffer (1M glycine, pH 3.0).

For IP experiments without cross-linking, 500 µg of nuclear extracts was pre-cleared with control IgG and subsequently incubated for 6 h at 4ºC with desired primary antibody. Protein A/G sepharose beads were added to the mixture and incubated for 2 h at 4ºC. The protein-
associated bead complexes were washed thrice with IP buffer and subsequently the eluates were probed with indicated antibodies.

For sequential IP experiments, nuclear extracts (1 mg) were immunoprecipitated first with 2 µg of α-SMAR1 antibody. Before proceeding for the second IP, minor fraction of the eluates was examined for the presence of HDAC6. Subsequently the eluate was immunoprecipitated with 1 µg of α-HDAC6. The final eluates were probed with α-Sam68 to evaluate the association.

**Chromatin immunoprecipitation (ChIP)**

ChIP assays were performed as mentioned in Luco et al [1] with some minor modifications. Approximately 2x10^6 HeLa cells were cross-linked with 1% formaldehyde for 10 min at room temperature. Cells were washed with ice cold PBS thrice and incubated in hypotonic buffer (25 mM HEPES pH 8.0; 1.2 mM MgCl_2; 10 mM KCl; 0.1% NP-40) containing protease inhibitor cocktail. Subsequently, the cells were homogenized and centrifuged at 3500 g for 5 min. The nuclei were collected and resuspended in sonication buffer (50 mM HEPES pH 8.0; 140 mM NaCl; 1mM EDTA; 0.1% Na-deoxycholate; 0.1% SDS) containing protease inhibitor cocktail. Nuclei were sonicated to yield DNA fragments of 200-500 bp length. Samples were centrifuged after addition of 1% Triton-X-100. To the supernatants, 2 µg of indicated antibodies were added and incubated for 8 h. To the immune-complexes, BSA and salmon sperm DNA-coated Protein A/G agarose beads were added and incubated for 4 h. The precipitated immune complexes were washed with low-salt, high-salt, LiCl and Tris-EDTA (TE) buffer. After reverse cross-linking and proteinase K treatment, the samples were processed to extract DNA. Immunoprecipitated DNA were analyzed with SyBR-green (Bio-Rad) by real-time qPCR (ABI step one plus qPCR). The primer sequences used for the amplification of immunoprecipitated CD44 fragments were obtained from Batsché et al., 2006, upon request.
RNA immunoprecipitation (RIP)

RIP was carried out as mentioned in Luco et al [1] with some minor modifications. 4x10^6 HeLa cells were cross-linked with 1% formaldehyde at room temperature for 15 minutes. Cells were washed with ice cold PBS in diethylpyrocarbonate (DEPC) water and lysed in lysis buffer (50 mM Tris-Cl pH 8.0; 100 mM NaCl; 5 mM MgCl_2; 0.5% NP-40) and centrifuged to pellet the nuclei. Nuclei were lysed in lysis buffer containing 140 mM NaCl, 1% Triton-X-100 and 0.1% Na-deoxycholate. After sonication and centrifugation, the supernatants were incubated with 3 µg of SMAR1 and control IgG antibodies for 8 h at 4°C. Protein A/G agarose beads that were previously blocked with BSA and yeast tRNA were added to the immune complexes and incubated for 4 h. Subsequently, the bead-bound complexes were washed with low-salt (150 mM NaCl) and high-salt (500 mM NaCl) RIPA buffer (50 mM Tris-Cl pH 8.0; 1mM EDTA; 1% Triton-X-100; 0.1% SDS; 0.1% Na-deoxycholate). Two washes with TE buffer were given before the beads were eluted with 0.1M NaHCO_3 and 200 mM NaCl. The eluates were reverse cross-linked and digested with 20 µg of proteinase K. RNA was isolated using liquid Trizol, reverse transcribed using M-MuLV reverse transcriptase (Invitrogen) and analyzed using SyBR-green for quantitative RT-PCR. The primer sequences used for the amplification of immunoprecipitated fragments were obtained from Batsché et al., 2006, upon request.

Modeling and docking studies

To determine the atomic details of interaction pattern of HDAC6-Sam68-SMAR1 complex, structure modeling of Sam68 and further protein-protein docking of HDAC6, SMAR1 and, Sam68 was performed. As the secondary structure information of K-rich domain (where the acetylation takes place in absence of HDAC6) is unavailable for Sam68, Prime module of Maestro version 9.6 was utilized for carrying out homology modeling of that specific domain of SAM68 protein [2]. The amino acids sequence for human Sam68 (NP_006550.1) was
obtained from swissprot database, K-rich domain consisting of 125 amino acids was considered for model building. PSI-BLAST template search was performed. The X-ray crystalline structure of star domain of quakin protein (4JVIH) [3] and the KH region of the Xenopus Star-GSG Quaking Protein (2BL5) [4] showed detectable degree of similarity with the query sequence, thus used as a template structures. The respective template coordinates were obtained from Protein Data Bank [5] and target sequence was aligned with the template structure. The co-ordinates of the conserved residues were used as the basis for modeling. The side chain coordinates for all non-identical residues were predicted using PRIME. Further, the model structure was refined for non-template regions. The structural coordinates were further optimized in the protein preparation wizard [6]. All the missing hydrogen atoms were added and further subjected to energy minimization with OPLS-2005. In addition SMAR1 structure was built with I-TASSER online server and zinc finger domain of HDAC6 (3C5K) structure [7] was optimized in the protein preparation wizard [6]. Interaction pattern of HDAC6-SMAR1-SAM68 with the U-rich region of CD44-v5 mRNA was analyzed. 3D structure for U-rich region of RNA was built in 'make-rna server' (http://casegroup.rutgers.edu/). The RNA sequences that was taken in to consideration, Seq: \(5' \text{A}^1 \text{G}^2 \text{U}^3 \text{G}^4 \text{C}^5 \text{U}^6 \text{C}^7 \) \(3' \text{U}^5 \text{U}^6 \text{A}^5 \text{A}^4 \text{A}^3 \text{A}^2 \text{G}^1 \). Both the RNA models were further docked in to the HDAC6-SMAR1-SAM68 complex using Z-Dock (v.3.0.2). The interaction models for SMAR1, K-rich Sam68, and HDAC6 were obtained with Z-Dock (v.3.0.2) [8]. ZDOCK predicts the interacting models by the fast Fourier transform; that allows 3D searches of spatial degrees of freedom between macromolecules. The predicted docked models of SMAR1-SAM68 and HDAC6-Sam68-SMAR1 were analyzed with the PyMol [9].

**Tail vein injections and histopathological evaluation**

MCF-7 cells were stably knockdown upon transduction with SMAR1-shRNA (sh3) lentivirus and control non-silencing (NS) virus. 1x10^6 stable cells were injected via tail vein in NOD-
SCID mice of 4-5 weeks. These mice were implanted with estradiol pellets of dosage 60 mg/60 days. Post 35 days of injections, mice injected with SMAR1-shRNA MCF-7 cells were observed to be lean and have lost body weight. The mice were sacrificed and the lungs were observed for metastatic tumorous nodules. Lung tissues were collected from the control and SMAR1 shRNA mice. These lung tissues containing tumor nodules were further processed for histological sectioning (5-10µm) as mentioned [10, 11]. The sections were stained with hematoxylin and eosin [12, 13] and observed under light microscopy. Immunohistochemical staining of the sections was done [14] with various antibodies of CD44 such as v3, v5, v6, v7 and v3-v10. All the antibodies were purchased from Abcam and the dilutions were used as recommended by the manufacturer. Sections were observed under bright-field microscope and photographed (Carl Zeiss, Germany).
**Fig. S1.** SMAR1 regulates alternative splicing.

(A) SMAR1 protein sequence analyzed by RNABindR software ([http://einstein.cs.iastate.edu/RNABindR/](http://einstein.cs.iastate.edu/RNABindR/)) for predicting the RNA-binding residues. Amino acid residues highlighted in red circle and represented with a positive signs are predicted to have RNA-binding affinity. Multiple arginine and serine residues along with proline are highlighted in red circles.

(B) *CD44* photinus luciferase v5 minigene was transfected along with renilla luciferase (transfection control), Flag-SMAR1 or SMAR1-sh745 for overexpression and knockdown of SMAR1, respectively. Inclusion of v5 alternate exon results in the expression of luciferase.
protein. Dual-Glo luciferase assays and further normalization of luciferase with renilla indicated that knockdown of SMAR1 enhanced CD44 v5 exon inclusion (as indicated in sh-SM bar). Error bars represent standard deviation (SD) from three independent experiments.

(C) Western blot analysis indicating the efficiency of SMAR1 overexpression with p3X-FLAG-SMAR1 (FSM) and knockdown (shSM) with sh745 plasmids (Sinha et al., 2010). FV is empty Flag-vector and shCntrl is scrambled shRNA vector.

(D) Semi quantitative-RT PCR analysis of CD44 variable exons (v7-v10) upon SMAR1 knockdown (left panel) and upon over-expression of SMAR1 (right panel).

(E) Semi quantitative-RT PCR analysis of CD44 variable exons (v7-v10) in indicated breast cancer cells (left panel) and upon ectopic expression of SMAR1 (right panel) in indicated breast cancer cell lines.

**Fig. S2.** ERK-dependent SMAR1 Expression. Western blot analysis to analyze the time-dependent expression of SMAR1 upon activation of ERK-1/2 using PMA (50 mg/ml). α-
Actin served as loading control, while the expression of phospho-p44/42 served as control for the activation of ERK-1/2.

**Fig. S3.** Occupancy of RNA pol II along the CD44 locus. (A-B) Chromatin immunoprecipitation (ChIP) experiments in formaldehyde-fixed HeLa cells upon treatment with PMA (50 ng/ml, 2 h) to check the occupancy of total RNA Pol II (A) and serine 2 phosphorylated RNA Pol II (B).
**Fig. S4.** Activation of ERK-MAPK pathway phosphorylates SMAR1 at threonines 345 and 360.

(A) In vivo phosphothreonine-proline (pTP) analysis by immunoblotting SMAR1 immunoprecipitates from HeLa nuclear and cytoplasmic fractions, upon treatment with U0126 (25 µM) for 30 min prior to PMA treatment.

(B) Western blot analysis of whole cell, nuclear and cytoplasmic fractions from HeLa cells transfected with wild type Flag-SMAR1 (FSM) and mutant Flag-SMAR1 [FSM (Mut)] followed by treatment with PMA (50 ng/ml) for 2 h. Differential localization of wild-type and mutant SMAR1 was analyzed by incubating with indicated antibodies.

(C) Sequence representation of KLH-conjugated SMAR1 peptides that were injected in rabbit to generate polyclonal antibodies.
(D) Western blot analysis of whole cell, nuclear and cytoplasmic fractions from HeLa cells that were either left untreated or treated with PMA (50 ng/ml) for indicated time points. Samples were immunoblotted with T345/T360-phospho and T345A/T360A-mutant antibodies.

Fig. S5. SMAR1 interacts with Sam68 and maintains it in deacetylated state.

(A) Immunoprecipitation of nuclear extracts using α-SMAR1 and α- Sam68. Eluates were immunoblotted with indicated antibodies. Endogenously SMAR1 is observed to interact with Sam68.
(B) RNase A treated (200 µg/ml) nuclear extracts were immunoprecipitated with α-SMAR1 and α-Sam68 antibodies and eluates were probed with indicated antibodies. Treatment with RNase A abolished the interaction.

(C-E) Docked pose of Sam68 (green) with SMAR1-interacting domain (blue spheres), sharing complimentary surfaces in anterior view (B) and posterior view (C). The interacting residues are closely highlighted in figure (D) showing SMAR1(green)-K-rich Sam68(gray) binding surface with more interactions. A major contribution in interactions is through the lysines of Sam68.

(F) Sam68 acetylation assays in HeLa cells treated with U0126 prior to PMA treatment.

(G) HeLa cells stably transduced with NS and SMAR1-sh3 lentivirus were assayed for Sam68 acetylation in the presence and absence of U0126.

(H) Representation of efficiency of Sam68 knockdown in MCF-7 cells that were stably knockdown using shSam68 lentivirus. ShSam68 clone 45 showed an efficient knockdown and was subsequently used to generate Sam68−/− MCF-7 cells.

(I-K) Semi-quantitative RT-PCR analysis of CD44 splice reporter minigene (I), FAS minigene (J) and Bcl-X (K) alternative splicing upon co-transfection along with full length and different truncations of myc-Sam68 to assess the splicing efficiency post ERK-MAPK pathway activation.

In, Input; IgG, Control IgG; IP, Immunoprecipitate. * Represents IgG.
Fig. S6. Interaction of SMAR1 and Sam68 with HDACs.

(A) HeLa cell nuclear extracts were immunoprecipitated with α-SMAR1 antibody and the eluates were immunoblotted for the presence of HDAC1 and Sam68. In, Input; IgG, Control IgG; IP, Immunoprecipitate.

(B) Immunoprecipitation (IP) assays with α-HDAC1 antibody and immunoblotting with indicated antibodies.

(C) IP experiments with α-Sam68 antibody in HeLa cell nuclear extracts and immunoblotting of the eluates with indicated antibodies.

(D-E) HeLa cell nuclear extracts were immunoprecipitated with α-SMAR1 antibody (D) and α-HDAC2 antibody (E). Eluates were immunoblotted with indicated antibodies for the presence of HDAC2 and SMAR1.

(F) RNA immunoprecipitation assays using α-HDAC6 and control α-IgG antibody in HeLa cell nuclear extracts.
(G) Western blotting analysis of different fractions obtained after the HeLa cell nuclear extracts are subjected to size-exclusion chromatography (Akta explorer; Superose 6, 10/300 GL). The fractions were blotted with indicated antibodies in the presence (RNase A +, lower panel) and absence of RNase A (RNase A -, upper panel), indicates the dissociation of HDAC6-SMAR1-Sam68 complex upon treatment with RNase.

(H) Acetylation assays by immunoprecipitation with α-acetyl lysine (Ac.K) antibody in the nuclear extracts of HeLa cells knocked down with HDAC1 siRNA. The eluates and nuclear extracts are blotted with indicated antibodies, highlighting no acetylation on Sam68 upon HDAC1 knockdown.

(I) Semi quantitative-RT PCR analysis of CD44 variable exons upon knockdown of HDAC1 in HeLa cells, indicates no change in alternative splicing.

* Represents IgG.

Supplementary Information, Table S1

<table>
<thead>
<tr>
<th>Primers Used for CD44 pET-v5-Luc PCR</th>
</tr>
</thead>
</table>
| N5Ins | GAGGGATCCGCTTCTTCCTGCC
| Luc 3' | CCAGCCGATAGAATGGCGCCG |
| Primers Used for CD44 Exon-Specific PCRs |
| hs5' | GATGGAGAAAGCTCTGAGCATC |
| hs3' | TTTGCTCCACCTTCTTTGACTCC |
| Human C13 | AAGACATCTACCCAGCAAC |
| Human e5FP | CATCCACAGACGAAGACAGTC |
| Human v2r | TGTGAAGATGATTCTTTGACTC |
| Human v3r | CATCATCAATGCCTGATCCAGA |
| Human pv4 | TCAACCACACCACGGGCTTT |
| Human pv5 | GTAGACAGAAATGGCACCAC |
| Human v6r | CAGCTGTCCCTGTGTGGCAGA |
| Human v7r | TCCTGCTTGATGACCTGTC |
Table S1. Various primers used for performing *CD44* Luc minigene PCR and CD44 exon-specific PCRs.

**Supplementary Information, Table S2**

<table>
<thead>
<tr>
<th>SMAR1 (Residue ID)</th>
<th>Sam68 (Residue ID)</th>
<th>Distance (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phe-260</td>
<td>Tyr-171</td>
<td>4.3</td>
</tr>
<tr>
<td>Arg-424</td>
<td>Lys-87</td>
<td>&lt;6.0</td>
</tr>
<tr>
<td>Arg-290</td>
<td>Lys-85</td>
<td>2.8</td>
</tr>
<tr>
<td>Arg-324</td>
<td>Lys-182</td>
<td>3.8</td>
</tr>
<tr>
<td>Tyr-343</td>
<td>Lys-183</td>
<td>2.8</td>
</tr>
<tr>
<td>Gln-426</td>
<td>Lys-99</td>
<td>5.1</td>
</tr>
</tbody>
</table>

Table S2. SMAR1 interacts with Sam68 and the following Hydrogen bond/salt bridge interactions are observed between SMAR1 and Sam68.

**Supplementary Information, Table S3**

<table>
<thead>
<tr>
<th>RNA sequence</th>
<th>SMAR1</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>C7(H41)</td>
<td>THR-504(O)</td>
<td>~2.4 Å</td>
</tr>
<tr>
<td>C7(H41)</td>
<td>LEU-505(O)</td>
<td>~2.6 Å</td>
</tr>
<tr>
<td>C7(HO2*)</td>
<td>GLN-506(O)</td>
<td>~1.8 Å</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RNA sequence</th>
<th>Sam68</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>C5(O2P)</td>
<td>MET-190(N)</td>
<td>~2.5 Å</td>
</tr>
<tr>
<td>C4(O1P)</td>
<td>ASP-191(N)</td>
<td>~2.9 Å</td>
</tr>
</tbody>
</table>

Table S3. Interacting residues among RNA and SMAR1-Sam68-HDAC6 binding
## Supplementary Information, Table S4

<table>
<thead>
<tr>
<th>Energy (kJ/mol)</th>
<th>HDAC6-SMAR1-Sam68</th>
<th>HDAC6-SMAR1-Sam68</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Energy</td>
<td>-151768.1</td>
<td>-141886.2</td>
</tr>
<tr>
<td>Stretch</td>
<td>7108.8</td>
<td>6994.6</td>
</tr>
<tr>
<td>Bend</td>
<td>7805.4</td>
<td>7251.4</td>
</tr>
<tr>
<td>Torsion</td>
<td>9583.9</td>
<td>8562.9</td>
</tr>
<tr>
<td>Improper Torsion</td>
<td>249.2</td>
<td>242.6</td>
</tr>
<tr>
<td>VDW</td>
<td>-7058.7</td>
<td>-6582.5</td>
</tr>
<tr>
<td>Electrostatic</td>
<td>-82397.3</td>
<td>-92588.7</td>
</tr>
<tr>
<td>Solvation</td>
<td>-87059.5</td>
<td>-65766.5</td>
</tr>
</tbody>
</table>

Table S4. Table representing the energy contents of the triplex complex HDAC6-SMAR1-Sam68 alone and in complex with RNA.

## SUPPLEMENTAL REFERENCES


