Cell

Phosphorylation-Regulated Binding of RNA Polymerase II to Fibrous Polymers of Low-Complexity Domains

Ilmin Kwon,[1,5] Masato Kato,[1,5] Siheng Xiang,[1] Leeju Wu,[1] Pano Theodoropoulos,[1] Hamid Mirzaei,[1] Tina Han,[1,3] Shanhai Xie,[1,4] Jeffry L. Corden,[2] and Steven L. McKnight[1,*]

[1]Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX 75390-9152, USA

[2]Department of Molecular Biology and Genetics, The Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

[3]Present address: Department of Physiology, Howard Hughes Medical Institute, University of California, San Francisco, San Francisco, CA 94158, USA

[4]Present address: Peloton Therapeutics, Dallas, TX 75235, USA

[5]These authors contributed equally to this work

*Correspondence: steven.mcknight@utsouthwestern.edu

http://dx.doi.org/10.1016/j.cell.2013.10.033

SUMMARY

The low-complexity (LC) domains of the products of the fused in sarcoma (FUS), Ewings sarcoma (EWS), and TAF15 genes are translocated onto a variety of different DNA-binding domains and thereby assist in driving the formation of cancerous cells. In the context of the translocated fusion proteins, these LC sequences function as transcriptional activation domains. Here, we show that polymeric fibers formed from these LC domains directly bind the C-terminal domain (CTD) of RNA polymerase II in a manner reversible by phosphorylation of the iterated, heptad repeats of the CTD. Mutational analysis indicates that the degree of binding between the CTD and the LC domain polymers correlates with the strength of transcriptional activation. These studies offer a simple means of conceptualizing how RNA polymerase II is recruited to active genes in its unphosphorylated state and released for elongation following phosphorylation of the CTD.

INTRODUCTION

Numerous forms of cancer result from translocation events wherein the amino terminal, low-complexity (LC) domains of any of three related RNA-binding proteins become fused to a variety of different DNA-binding domains (Arvand and Denny, 2001; Guipaud et al., 2006; Lessnick and Ladanyi, 2012). The relevant RNA-binding proteins include the products of the fused in sarcoma (FUS), Ewings sarcoma (EWS), and TAF15 genes. Collectively, these three RNA-binding proteins are referred to as FET (FUS/EWS/TAF15) proteins (Andersson et al., 2008). The proteins encoded by these varied translocation products are causative of transformation (Bertolotti et al., 1999; Crozat et al., 1993; Ichikawa et al., 1999; Rabbitts et al., 1993; Zinszner et al., 1994). They further display dual dependency upon both the DNA-binding domain, which can be represented by members of the homeobox, zinc finger, ETS, or leucine zipper families of DNA-binding domains, as well as the LC domains donated by FUS, EWS, or TAF15. The DNA-binding domains are understood to direct the cancer-causing fusion proteins to appropriate batteries of genes adequate to facilitate cell growth or survival. By contrast, the LC sequences donated by members of the FET family are understood to function as transcriptional activation domains.

Over the past several decades the fields of biochemistry, biophysics, and molecular biology have achieved a concrete understanding of how DNA-binding domains function. The atomic structures of many such domains have been resolved by either X-ray crystallography or NMR spectroscopy, often in complex with their specific DNA substrates. By contrast, far less is known about the manner in which transcriptional activation domains operate at a mechanistic level. Prototypic activation domains are comprised of low-complexity sequences that exist in an intrinsically disordered, random coil conformation (Huntley and Golding, 2002; Uversky, 2002). The observation that evolutionary pressure of cancer cell formation has repeatedly led to the selection of the LC domains of FET proteins as the fusion partner to a variety of different DNA-binding domains strongly hints that these particular LC sequences may be exceptionally potent transcriptional activation domains.

The LC domain of FUS donated to the translocation product causative of cancer has a highly skewed distribution of amino acids. Of the 220 residues within the FUS LC domain, 84% are comprised of only four amino acids—glycine, serine, glutamine, and tyrosine. The domain contains zero representatives of glutamic acid, lysine, arginine, cysteine, histidine, valine, leucine, isoleucine, tryptophan, or phenylalanine. By having an amino acid composition dominated by only a four letter code (G, S, Q, and Y), the LC domain of FUS would appear to be more like nucleic acids than typical proteins that fold into their ultimate, three-dimensional shape by use of a much wider reliance on all 20 types of amino acid residues. When incubated at high concentrations, the LC domain of FUS has been found to polymerize

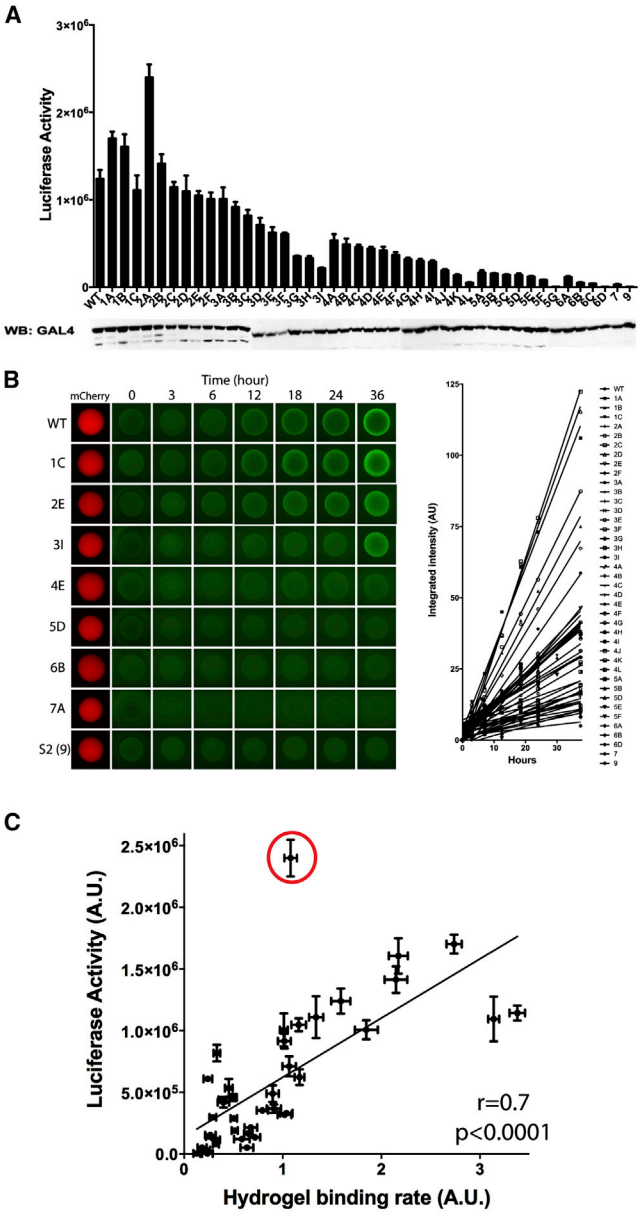

Figure 1. Correlation Between Transcriptional Activation and Hydrogel Binding of Native and Mutated Derivatives of the LC Domain of FUS

(A) The LC domain of FUS was linked to the DNA-binding domain of GAL4 and assayed for activation of a GAL4 reporter gene in transiently transfected U20S cells. Activity of the native LC sequence was compared with 43 variants wherein different number of tyrosine residues were randomly mutated to serine (see text). Identities of specific tyrosine-to-serine changes in each mutant are shown in Table S1. Expression levels for all test protein were assayed by western blotting as shown below histograms.

(B) The GFP-linked LC domains of FUS carrying the same mutations as (A) were exposed to mCherry:FUS hydrogels (left) and initial binding rates were measured (right).

(C) A correlation plot between the transactivation activity and hydrogel-binding rate of the individual LC mutants. Note that there is but one significant outlier indicated by a red circle. This is the "2A mutant" and described in more detail in the text and Figure 5.

into uniform, amyloid-like fibers (Han et al., 2012; Kato et al., 2012). Although morphologically similar to pathogenic amyloid aggregates, polymeric fibers formed from the LC domain of FUS are labile to depolymerization, raising the possibility that reversible LC polymerization may engender functional utility. Given two measurable features of the LC domain of FET proteins (transcriptional activation capacity and polymerization propensity), we hereby describe experiments that test the correlative relationship between the two.

RESULTS

The FUS LC domain contains 27 repeats of the triplet sequence [G/S]Y[G/S], and derivatives carrying 5, 9, 15, and 27 tyrosine-to-serine mutations in randomly chosen triplets showed progressively diminished capabilities of polymerization and stress granule recruitment (Kato et al., 2012). In order to initiate studies of the transcriptional activation capacity of the FUS LC domain, it was linked to the GAL4 DNA-binding domain and assayed by transient transfection for activation of firefly luciferase activity driven by a GAL4-dependent reporter gene. Forty-three mutated derivatives were prepared wherein individual tyrosine residues of the triplet sequence within the LC domain were randomly mutated to serine. Three mutants randomly changed a single tyrosine, 6 changed two tyrosines, 9 changed three tyrosines, 12 changed four tyrosines, 7 changed five tyrosines, 4 changed six tyrosines, 1 changed seven tyrosines, and 1 changed nine tyrosines (Table S1). As shown in Figure 1A, concordance was observed between the number of tyrosine residues changed to serine and the degree of impediment upon of transcriptional activation. With a single exception, variants carrying either one or two tyrosine-to-serine changes exhibited transcriptional activation capacity indistinguishable from the GAL4:wild-type fusion protein. The exceptional variant (2A) was considerably more active than the wild-type LC domain of FUS. Variants carrying three tyrosine-to-serine changes tended to be slightly less active than those carrying zero, one or two mutations. The trend in loss of activity was observed to track with variants carrying four, five, or six random tyrosine-to-serine mutations, and mutants carrying seven or nine tyrosine-to-serine mutations revealed no detectable capacity to activate expression of the GAL4-responsive luciferase target gene.

A simple assay for the propensity of LC domains to polymerized into amyloid-like fibers has been described previously (Kato et al., 2012). Briefly, upon prolonged incubation of a hybrid protein linking the LC domain of FUS to mCherry, the protein polymerizes into a hydrogel-like state. Microscopic gel droplets can be formed in chamber slides, exposed to soluble test protein, then scored for trapping of the test protein via confocal microscopy. The molecular basis of trapping has been deduced, by TIRF microscopy, to result from copolymerization of the GFP-tagged test protein into existing polymeric fibers in the mCherry:FUS hydrogel droplets (Kato et al., 2012). Forty out of the 43 variants described above were fused to GFP and tested for their ability to be trapped by hydrogel droplets composed of mCherry linked to the wild-type LC domain of FUS. As shown in Figure 1B, variants carrying one or two tyrosine-to-serine changes were trapped by the mCherry:FUS hydrogel droplets in a manner

similar to the protein linking the wild-type FUS LC domain to GFP. Sequential diminishment in hydrogel trapping was observed for variants carrying three, four, five, six, seven, or nine tyrosine-to-serine mutations. The scatter plot shown in Figure 1C directly compares the transcriptional activation and hydrogel trapping capacities of all 40 mutants that were scored in both assays. With the exception of the 2A mutant, which activates transcription considerably better than the wild-type LC domain of FUS, a strong correlation was observed between the two activities under study (Pearson's r = 0.7 with a p value of less than 0.0001). These data tentatively indicate that transcriptional activation capacity of the LC domain of FUS may be dependent upon its ability to polymerize.

B-isox Microcrystals Precipitate Nuclear Proteins

Having observed a correlative relationship between fiberization and transcriptional activation capacity, we prepared nuclear extracts from HEK293 cells and subjected the lysate to precipitation by the biotinylated isoxazole (b-isox) chemical. As described previously, the b-isox chemical crystallizes in cold aqueous buffer at concentrations between 30 and 100 μM (Kato et al., 2012). X-ray diffraction studies at high resolution revealed the b-isox crystal surface to contain a wavy repetition of peaks and troughs separated by 4.7 Å, and it has been hypothesized that the troughs are properly disposed to capture proteins containing significant stretches of random coil polypeptide. These disordered, LC sequences are thought to adopt an extended β strand conformation having dimensions that match the trough of the b-isox crystal surface, thereby facilitating selective precipitation of proteins endowed with substantive regions of disordered polypeptide.

Given that earlier studies were limited to cytoplasmic lysates, we prepared nuclear lysates from cultured mammalian cells, precipitated with b-isox microcrystals, separated the precipitated proteins on SDS polyacrylamide gels, and identified precipitated proteins by shotgun mass spectrometry (Experimental Procedures). Roughly 580 nuclear proteins were identified as being precipitated by b-isox microcrystals (Table S2). This survey revealed overlap with many of the proteins precipitated from cytoplasmic extracts by the b-isox chemical, including many hnRNP proteins. Additionally, however, these efforts revealed the identities of a number of nuclear proteins not observed in earlier experiments that had interrogated cytoplasmic lysates. Prominent among the newly identified, b-isox precipitated nuclear proteins were TAF15, the largest subunit of RNA polymerase II, numerous subunits of the mediator complex (including Med4, Med6, Med10, Med13L, Med 14, Med 15, Med 17, and all subunits of CDK8), a variety of enzymes involved in epigenetic modification of histones, both coactivator and corepressor proteins known to control the function of nuclear hormone receptors, and many distinct SR proteins known to be involved in pre-mRNA splicing. Parallel studies were performed on nuclear extracts prepared from *Saccharomyces cerevisae*, leading to the identification of roughly 260 yeast nuclear proteins significantly precipitated by the b-isox chemical (Table S3). We again observed b-isox-mediated precipitation of the largest subunit of RNA polymerase II. We also observed the list of precipitated proteins to contain multiple subunits of the mediator complex, and a number of gene-specific transcription factors endowed with glutamine- and asparagine-rich, low-complexity sequences.

With respect to nuclear proteins precipitated by the b-isox chemical, our initial focus fell on TAF15 and the largest subunit of RNA polymerase II. Among the 580 mammalian proteins selectively precipitated by b-isox microcrystals, TAF15 registered the 23rd highest number of spectral counts, and the largest subunit of RNA polymerase II registered the 46th highest number of spectral counts (Table S2). B-isox precipitation of yeast nuclear extracts also revealed the largest subunit of RNA polymerase II as displaying an exceptionally high number of spectral counts (Table S3).

TAF15 represents the third paralog of the FET family of RNA-binding proteins and is endowed with an N-terminal LC domain very similar to the LC domains found in FUS and EWS. Unlike FUS and EWS, TAF15 has been identified as a substoichometric component of the TFIID complex, which has been extensively characterized as an important assembly in transcriptional initiation by RNA polymerase II (Bertolotti et al., 1999). Attention to the largest subunit of RNA polymerase II was prompted by the fact that it contains an intrinsically disordered region of roughly 380 residues at its C terminus. This C-terminal domain (CTD) of mammalian RNA polymerase II contains 52 heptad repeats of the sequence $Y_1S_2P_3T_4S_5P_6S_7$ (Corden et al., 1985). Extensive work over the past two decades has confirmed the importance of the CTD in the transcription cycle (Buratowski, 2009; Egloff et al., 2012; Phatnani and Greenleaf, 2006). Included among a variety of seminal advancements, the field now understands that in its initial state of recruitment to a promoter, the CTD exists in the unphosphorylated state. A series of obligatory phosphorylation events on serine residues S_2, S_5 and S_7 of the CTD heptad repeats accompanies release of RNA polymerase from the preinitiation complex so that it can begin transcriptional elongation (Egloff and Murphy, 2008). When aligned contiguously, the CTD contains 52 repeats of the triplet sequence $S_7Y_1S_2$ (or close variations thereof). These repeats form a subset of the [G/S]Y[G/S] repeats that we have found to be critical for polymerization of the LC domain of FUS (Kato et al., 2012).

Phosphorylation of the CTD Blocks B-isox Precipitation

Using antibodies specific to either the unphosphorylated form of the CTD of RNA polymerase II (8WG16), or antibodies to S_2 phosphorylation (3E10), S_5 phosphorylation (3E8), or S_7 phosphorylation (4E12) (Chapman et al., 2007), we observed that b-isox microcrystals preferentially precipitate only the unphosphorylated form of the yeast CTD (Figure 2). Parallel analyses of b-isox precipitates from mammalian cells also revealed selective precipitation of only the unphosphorylated form of the largest subunit of RNA polymerase II (Figure S1A available online). We tentatively hypothesize that the CTD is the determinant of the largest subunit of RNA polymerase II that facilitates b-isox precipitation, and that phosphorylation on residues S_2, S_5 or S_7 somehow impedes the ability of the CTD to fit within the 4.7 Å surface troughs of b-isox microcrystals.

In order to test these ideas more carefully, a GFP fusion protein was prepared linked to heptad repeats 27-52 of the human CTD (Experimental Procedures). The purified, recombinant

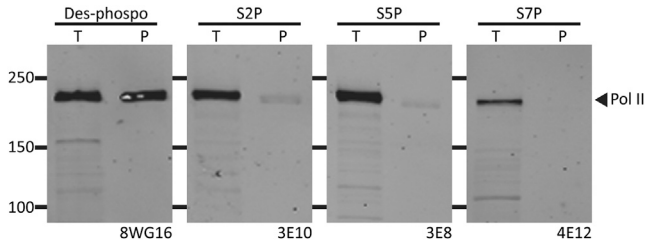

Figure 2. Biotinylated Isoxazole Preferentially Precipitates Unphosphorylated CTD of RNA polymerase II

Western blot of total cell extract (T) or b-isox pellet (P) with different CTD antibodies. Bands shown in the four panels correspond in size to the largest subunit of RNA polymerase II (217 kD). The 8WG16 antibody recognizes the unphosphorylated CTD, the 3E10 antibody recognizes serine 2 phosphorylated CTD, the 3E8 antibody recognizes serine 5 phosphorylated CTD, and the 4E12 antibody recognizes serine 7 phosphorylated CTD. B-isox significantly precipitates only unphosphorylated CTD of Pol II. See also Figure S1.

GFP:CTD_{C26} fusion protein was effectively precipitated by b-isox microcrystals, yet phosphorylation of the CTD by either CDK7, which phosphorylates S_5 and S_7 (Akhtar et al., 2009; Kim et al., 2009; Rickert et al., 1999; Roy et al., 1994; Trigon et al., 1998), or CDK9, which phosphorylates S_2, S_5 and S_7 (Czudnochowski et al., 2012; Jones et al., 2004; Ramanathan et al., 2001), prevented b-isox-mediated precipitation (Figure S1B). These observations confirm the CTD as the determinant of the largest subunit of RNA polymerase II responsible for b-isox precipitation and give evidence that when the CTD is phosphorylated, precipitation is prevented.

Binding of the CTD to Hydrogel Droplets Composed of FET Protein Polymers

Having recognized that the iterated $S_7Y_1S_2$ triplet repeats of the CTD correspond to a variant of the [G/S]Y[G/S] triplet repeats diagnostic of the LC domains of FET proteins, we tested whether the GFP:CTD fusion protein might bind to hydrogel droplets formed from the LC domains of the three FET proteins (FUS, EWS, and TAF15), as well as hydrogel droplets formed from the LC domains of hnRNPA2 and cold inducible RNA-binding protein (CIRBP). As shown in Figure 3A, the GFP:CTD fusion protein was avidly trapped by hydrogel droplets composed of the LC domain of TAF15. No trapping whatsoever was observed when the GFP:CTD test protein was exposed to either hnRNPA2 or CIRBP hydrogel droplets and weak trapping was observed for FUS and EWS hydrogels.

GFP:CTD binding to TAF15 hydrogel droplets was extended in two ways. First, we compared the binding properties of two GFP test proteins, one containing 20 heptad repeats from the highly conserved, N-terminal half of the CTD (repeats 1-20 = GFP:CTD_{N20}) and another containing 20 heptad repeats from the degenerate C-terminal half of the CTD (repeats 33-52 = GFP:CTD_{C20}). As shown in Figure 3B, the GFP test protein appended to the 20 degenerate CTD repeats, GFP:CTD_{C20}, bound TAF15 considerably more avidly than the counterpart fusion containing the 20 conserved CTD repeats (GFP:CTD_{N20}). Second, we compared hydrogel trapping for fusion proteins containing 5, 10, 15, 20, or all 26 of the degenerate CTD repeats (GFP:CTD_{C5} = CTD repeats 48-52, GFP:CTD_{C10} = CTD repeats 43-52, GFP:CTD_{C15} = CTD repeats 38-52, GFP:CTD_{C20} = CTD repeats 33-52, and GFP:CTD_{C26} = CTD repeats 27-52). As shown in Figure 3C, no trapping by mCherry:TAF15 hydrogel droplets was observed for the fusion protein containing only 5 heptad repeats, intermediate trapping was observed for the fusion containing 10 heptad repeats, and equivalent trapping was observed for fusions containing 15, 20, or 26 repeats.

When evaluating the kinetics of binding of the various GFP:CTD constructs to hydrogel droplets composed of mCherry linked to the LC domains of TAF15, FUS, or EWS, it was noticed that gel trapping occurred more rapidly than the homo- or heterotypic trapping of GFP fusion proteins linked to the LC domains of FET proteins. Knowing that the latter trapping occurs via a mechanism of copolymerization of soluble test proteins to fibrous polymers of which the hydrogel droplets are themselves composed (Kato et al., 2012), we employed fluorescence microscopy to evaluate binding of the GFP:CTD_{C26} fusion protein to fibrous preparations of mCherry:TAF15. As soon as the mixture of the two protein samples could be applied to the microscope substrate, clear evidence of GFP signal could be seen to coat the entire length of mCherry:TAF15 polymeric fibers (Figure S2). Instead of slow, time-dependent copolymerization extending from either end of existing fibers, the GFP:CTD_{C26} protein rapidly bound to the lateral surface of mCherry:TAF15 fibers. These observations offer evidence of a second and distinct means by which LC polymeric fibers are able to trap otherwise soluble test proteins composed, in part, of LC domains. They may further offer mechanistic insight into the perplexing promiscuity of protein:protein interaction of intrinsically disordered proteins (Cumberworth et al., 2013).

CTD Phosphorylation Blocks Hydrogel Binding

Knowing that phosphorylation of the CTD prevents precipitation by b-isox microcrystals, we evaluated the effects of phosphorylation on trapping of GFP:CTD test protein by TAF15 hydrogels. When phosphorylated by either CDK7 or CDK9, the GFP:CTD fusion protein was fully blocked from binding to hydrogels composed of polymers of the LC domain of TAF15 (Figure S3). Emboldened by this observation, we prebound the GFP:CTD_{C26} test protein to mCherry:TAF15 hydrogel droplets then added either CDK7 or CDK9 in the presence of ATP. In both cases, the GFP signal was lost in an enzyme concentration-dependent manner (Figure 4A). When applied at fixed enzyme levels, the GFP signal diminished in a time and ATP-dependent manner (Figure 4B). A video showing the dynamics of GFP signal loss can be seen in Movie S1. Material released from mCherry:TAF15 gel droplets in response to CDK7 was recovered from the wells of chamber slides and evaluated by western blotting using antibodies to both GFP and the S_5 phospho-form of the CTD. As shown in Figure 4C, both GFP and phosphorylated CTD S_5 western blotting signals were released from the mCherry:TAF15 hydrogels as a function of enzyme concentration. We conclude from these observations that not only does phosphorylation of the CTD prevent binding to mCherry:TAF15 hydrogels, but prebound material is accessible to kinase-mediated phosphorylation and subsequent release from the gels.

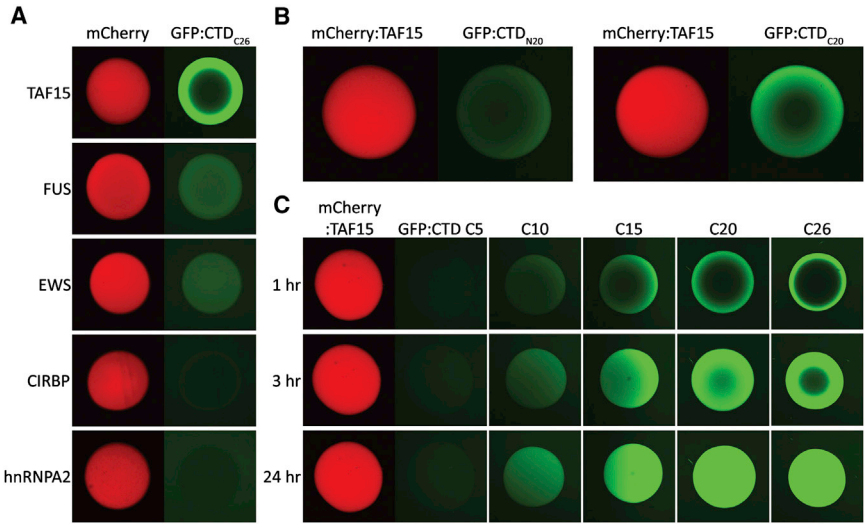

Figure 3. Selective Binding of GFP:CTD to mCherry:TAF15 Hydrogel Droplets

(A) Hydrogels composed of LC domains of mCherry:TAF15, mCherry:FUS, mCherry:EWS, mCherry:hnRNPA2 and mCherry:CIRP were incubated with a soluble form of GFP linked to the C-terminal 26 heptad repeats of the CTD of mammalian RNA polymerase II. Little or no retention of the GFP:CTD$_{C26}$ protein was observed for the mCherry:hnRNPA2 or mCherry:CIRBP hydrogel droplets, and weak binding was observed for mCherry:FUS and mCherry:EWS hydrogels. By contrast, strong retention was observed for mCherry:TAF15 hydrogel droplets. See also Figure S2.

(B) The degenerate C-terminal half of the CTD (GFP:CTD$_{C20}$) binds mCherry:TAF15 hydrogel droplets more strongly than the highly conserved, N-terminal half of the CTD (GFP:CTD$_{N20}$).

(C) The binding intensity of the GFP:CTD degenerate repeats to mCherry:TAF15 hydrogel droplets correlates with the number of heptad repeats, with no binding observed for GFP:CTD$_{C5}$, weak binding for GFP:CTD$_{C10}$ and strong binding for GFP:CTD$_{C15}$, GFP:CTD$_{C20}$ and GFP:CTD$_{C26}$.

The 2A Mutant of the FUS LC Domain Displays a Gain-of-Function Phenotype

In comparing the properties of mutants of the LC domain of FUS with respect to both fiber polymerization and transcriptional activation (Figure 1), we observed significant discordance for only one mutant. The so-called "2A mutant" bound to mCherry:FUS hydrogel droplets with avidity indistinguishable from the wild-type protein, yet activated transcription considerably better than the normal LC domain of FUS. Hydrogel droplets were formed from a fusion protein linking the 2A mutant to mCherry, fully analogous to the wild-type protein linked to mCherry save for the two tyrosine-to-serine changes. When challenged with the different heptad repeats of GFP:CTD test proteins, significantly stronger gel trapping was observed for hydrogel droplets formed from the 2A mutant LC domain (Figure 5A). We conclude that polymeric fibers formed from the 2A mutant bind the unphosphorylated CTD better than the wild-type LC domain of FUS, and offer the hypothesis that this gain-of-function phenotype may explain why the 2A mutant activates transcription better than the wild-type LC domain of FUS (Figure 1). Despite the enhanced binding of the 2A mutant to various GFP:CTD test proteins, the latter interactions were blocked by CDK7- or CDK9-mediated phosphorylation (Figure S4A). Moreover, application of either of these protein kinase enzymes, in the presence of ATP, caused prebound GFP:CTD$_{C26}$ protein to be released from hydrogel droplets composed of LC domain prepared from the 2A mutant in a manner indistinguishable of CTD release from hydrogel droplets formed from the wild-type LC domain of FUS (Figure S4B).

An alignment of the 27 [G/S]Y[G/S] triplet repeats located within the LC domains of FUS and TAF15 is shown in Figure 5B. The two tyrosine residues mutated to serine in the 2A, gain-of-function variant, correspond to tyrosine residues embedded within the 16th and 20th [G/S]Y[G/S] triplets of the FUS LC domain. In comparing this region of the two LC domains, FUS that when polymerized binds the CTD weakly, and TAF15 that binds the CTD avidly, we notice that triplet repeats 16-21 of the TAF15 LC domain are unusual. Instead of the central tyrosine residues being flanked by either glycine or serine, six straight repeats bear an aspartic acid residue on one side of the central tyrosine or the other (SYD$_{16}$, DYG$_{17}$, SYD$_{18}$, GYD$_{19}$, SYD$_{20}$, and NYD$_{21}$). As will be discussed, this "16-21 window" may be worthy of close attention with respect to the ability of the LC domains of FET proteins to capture the CTD of RNA polymerase II.

DNA-Mediated Enhancement of FET Protein Polymerization

The test tube polymerization of the LC sequences associated with FET proteins is only observed at high concentrations (Kato et al., 2012). One means by which appropriately high concentrations necessary for polymerization could take place in living cells is if FET proteins are iteratively bound to a multivalent substrate. ChIP-chip experiments have given evidence that when the LC domain of EWS is linked to the FLI DNA-binding domain, it binds in living cells to microsatellites carrying iterative repeats of the tetranucleotide sequence GGAA (Gangwal et al., 2008). Since GGAA represents an optimal binding site for the ETS DNA-binding domain of the EWS:FLI protein, the fusion protein binds iteratively to these otherwise inert microsatellites. If located within 5-10 kb of a target promoter, the translocation product-bound microsatellite is envisioned to become a rogue enhancer. Under such a setting, it is possible that the LC domains of the densely packed EWS:FLI molecules can be prompted to polymerize.

As a test of this hypothesis, microsatellite sequences associated with the *hNR0B1* gene were amplified by PCR from U2OS cells (Experimental Procedures). A PCR product containing 25 repeats of the GGAA tetranucleotide was incubated with a recombinant protein linking mCherry to a fusion protein composed of the DNA-binding domain of FLI attached to the LC domain of FUS (Experimental Procedures). In the absence of added DNA, the 0.5 μM levels of the mCherry:FUS-FLI protein

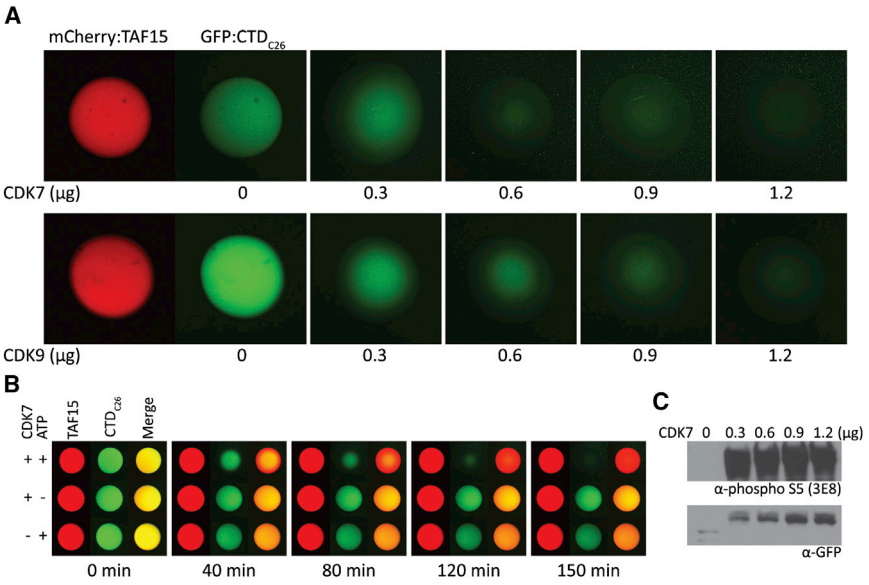

Figure 4. CDK7-Mediated Release of GFP:CTD$_{C26}$ from mCherry:TAF15 Hydrogel Droplets

(A) The GFP:CTD$_{C26}$ trapped by mCherry:TAF15 hydrogel was released upon addition of CDK7 or CDK9 in the presence of ATP in an enzyme concentration-dependent manner.

(B) Hydrogel droplets of mCherry:TAF15 (red) were exposed to GFP:CTD$_{C26}$ (green). Chamber slides containing individual hydrogel droplets were exposed to CDK7 plus ATP (top row), CDK7 alone (middle row), or ATP alone (bottom row). Samples exposed to both CDK7 and ATP revealed the time-dependent release of GFP:CTD$_{C26}$. See also Movie S1.

(C) Western blot analysis of materials released into the well of chamber slides revealed progressive increases in the presence of soluble, phosphorylated CTD (top), and soluble GFP:CTD$_{C26}$ (bottom).

See also Figure S3.

remained in a largely disaggregated state as deduced by transmission electron microscopy (Figure 6B). By contrast, addition of the PCR product containing the iterative GGAA repeats led to the formation of an extensive network of interwoven fibers (Figure 6A). When evaluated by a gel mobility shift assay, it was apparent that multiple copies of the mCherry:FUS-FLI protein were simultaneously bound to the probe DNA containing iterative repeats of the GGAA tetranucleotide (Figure S5). We tentatively conclude that when multiple copies of the mCherry:FUS-FLI protein are simultaneously bound to DNA, polymerization of the LC domain of FUS is enhanced.

Polymerization-Dependent CTD Binding to the LC Domain of TAF15

Hydrogel droplets formed from a fusion protein linking mCherry to the amino terminal 208 residues of TAF15 bind a variety of GFP:CTD constructs, particularly those housing 15 or more of the degenerate, C-terminal heptad repeats (Figure 3C). A truncated version of the TAF15 LC domain restricted to 80 amino-terminal residues was observed to form hydrogel droplets equally as well as the intact LC domain, yet was only minimally capable of trapping the GFP:CTD$_{C10}$ and GFP:CTD$_{C20}$ constructs (Figure S6). Addition of the remaining 128 residues of the LC domain yielded a protein that readily formed hydrogels that trapped the GFP:CTD constructs avidly.

These observations hint to the possibility that the amino terminal half of the TAF15 LC domain may be critical for polymerization, and that the carboxyl terminal half might be critical for CTD binding. Proceeding with this idea, a series of 48 variants was prepared wherein between one and six tyrosine residues of the [GS]Y[GS] triplet repeats were randomly mutated to serine (Table S4). Each mutant was linked to the DNA-binding domain of GAL4 and tested for its ability to activate the GAL4-luciferase reporter gene. As shown in Figure 7A, all six variants carrying a single tyrosine-to-serine mutation were less competent in transcription activation capacity than the wild-type protein. Variants carrying two, three, four, five, or six random tyrosine-to-serine mutations suffered sequentially diminished activity. Three of the 48 mutants were chosen for further study. These include the 1F, 2H, and 3K mutants, all of which were among the most debilitated variants within their respective classes. All of the tyrosine residues altered in these particularly deleterious mutants localized to the first 80 residues of the TAF15 LC domain that, on its own, is capable of fiberization and hydrogel formation. The 1F mutant changes tyrosine 46 to serine, the 2H mutant changes tyrosines 38 and 56 to serine, and 3K changes tyrosines 17, 46, and 63 to serine (Table S4).

The TAF15 LC domains of these variants were linked to mCherry and compared to the wild-type LC domain with respect to hydrogel formation, fiberization, and CTD interaction. None of the mutants formed hydrogels when incubated under standard conditions (Movie S2). When incubated under conditions favorable for fiberization, the mCherry variant linked to the wild-type LC domain of TAF15, as visualized by fluorescence microscopy, formed obvious tangles of aggregated fibers. Almost no polymerization was observed for the 1F, 2H, and 3K mutants. Inspection of the rare occurrence of visible mCherry aggregates in the latter variants revealed small amorphous particles (Figure 7B). Finally, when mixed with the GFP:CTD$_{C26}$ construct and subjected to coimmunoprecipitation, no interaction whatsoever was observed for anything but the mCherry fusion protein linked to the native LC domain of TAF15 (Figure 7C). These and other data presented herein give evidence that the LC domains of FET proteins obligatorily rely upon polymerization in order to execute their role in transcriptional activation. Whether LC domain polymerization might be more broadly employed in the context of gene regulation in eukaryotic cells remains open to question.

Schematic Concept of RNA Polymerase Recruitment by LC Domain Polymers

Cancer cells expressing the translocation product wherein the LC domain of EWS is linked to the ETS DNA-binding domain of

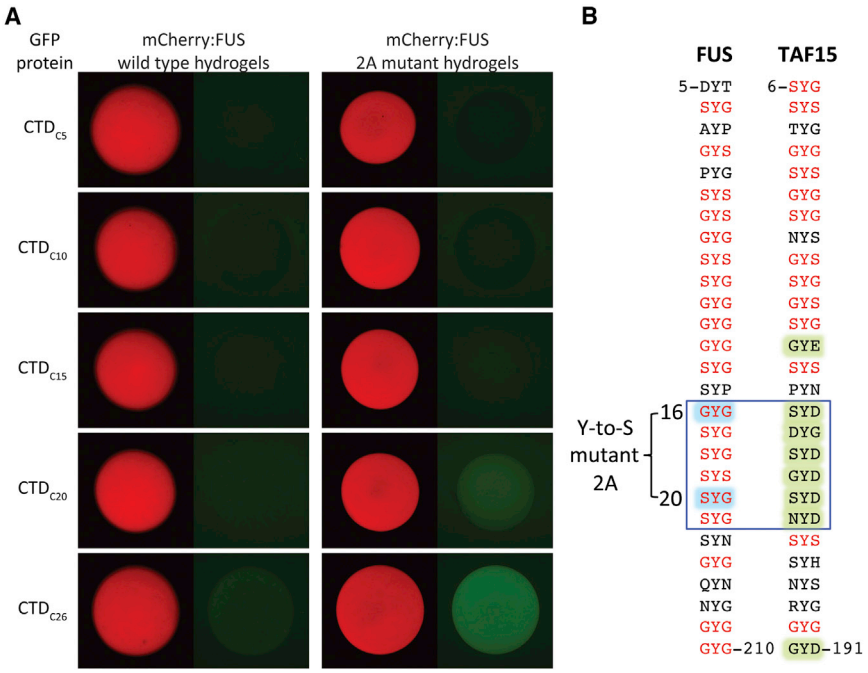

Figure 5. Hydrogels Formed from the 2A Mutant of FUS LC Domain Display Enhanced CTD Binding

(A) 2A mutant of FUS LC domain, which exhibits enhanced transactivation activity relative to the wild-type protein (Figure 1C), carries tyrosine-to-serine mutations at tyrosine within the 16th and 20th [G/S]Y[G/S] triplet repeats (Figure 5B). The 2A mutant LC domain was fused with mCherry, and hydrogel-binding assays were carried out with GFP:CTD fusions carrying different numbers of heptad repeats. Compared with hydrogel droplets formed with the native LC domain of FUS, those formed from the 2A mutant showed enhanced binding to GFP:CTD$_{C20}$ and GFP:CTD$_{C26}$. See also Figure S4.

(B) Alignment of [G/S]Y[G/S] triplet repeats in low-complexity domains of FUS (left) and TAF15 (right). The two tyrosine-to-serine mutations causing FUS to suffer a gain-of-function enhancement in transcriptional activation and CTD binding are highlighted in blue. All triplet repeats of TAF15 carrying an aspartic acid or glutamic acid adjacent to tyrosine are highlighted in green. Canonical [G/S]Y[G/S] triplet repeats of both LC domains are shown in red.

FLI depend upon this fusion protein to drive cell growth or survival. The EWS:FLI translocation product is undoubtedly employed as an aberrant transcription factor required to reprogram gene expression in a manner beneficial to cancer cells. We hypothesize that upon binding to the relevant sites on DNA, including microsatellites bearing iterative repeats of the tetranucleotide sequence GGAA, the EWS:FLI fusion protein can polymerize via its LC domains. As depicted schematically in (D) of Figure 7, we further hypothesize that fibrous polymers of the LC domain create binding sites for the CTD of RNA polymerase. Finally, we predict that phosphorylation of the CTD by CDK7, CDK9, or a related protein kinase enzyme facilitates release of the polymer-bound RNA polymerase so that it can escape gene promoters and enter into the process of transcriptional elongation.

DISCUSSION

Here, we report evidence that the LC domains of FET proteins may require polymerization as a means of directly capturing the CTD of RNA polymerase II. If correct, these data lend credence to the hypothesis advanced by Kato, Han, and colleagues (Han et al., 2012; Kato et al., 2012). The thesis of the initial studies on b-isox precipitation, LC sequences, LC polymerization, and LC-derived hydrogels offered the concept that intrinsically disordered domains associated with RNA regulatory proteins and transcription factors might reversibly polymerize as a means of helping establish cellular organization and information flow from DNA to RNA to protein. The present study extends earlier work in several ways. First, by having a simple functional read-out (transcriptional activation), we have been able to establish a pattern of significant concordance between the capacity of LC domains to polymerize and their ability to function as transcriptional activation domains (Figure 1). Indeed, the single most discordant mutant (the 2A mutant) was found to suffer a gain-of-function phenotype in transcriptional activation capacity that may be attributable to its enhanced capacity to bind the CTD of RNA polymerase II. Second, we offer the simple conclusion that, when appended to a DNA-binding domain as a function of scores of independent cancer-causing translocation events, the LC domains of FET proteins act to directly recruit the CTD of RNA polymerase II (Figure 3). In order to achieve this task, the LC domains of FET proteins must be capable of polymerization (Figure 7). Whereas the hydrogel droplets deployed as our mainstay assay are composed of long fiberous polymers containing thousands of subunits, we have no idea as to how many subunits must polymerize as a "point source" to recruit the CTD of RNA polymerase. It is entirely possible that fiberous seeds composed of only a handful of LC domain subunits might be sufficient to nucleate functionally competent organizational puncta.

Evidence that the LC domains of FET proteins are able to interact with the CTD of RNA polymerase II is consistent with recent studies of FUS in its native form (Schwartz et al., 2012). The latter study provided compelling evidence for direct interaction between the CTD of RNA polymerase II and the intact FUS protein. Intriguingly, this interaction was shown to be RNA dependent. In a new study, Cech and colleagues offer exciting data that may explain RNA dependence (Schwartz et al., 2013). Since the intact FUS protein is endowed with an RNA-binding domain, addition of an RNA polymer offers the opportunity for FUS to bind multiple times to the RNA substrate, thereby bringing LC domains of multiple protomers into proximity, and thereby enhancing local concentration and the propensity of the LC domains to polymerize.

Similar phenomena may well be in play in the context of translocation products carrying the LC domains of FET proteins fused

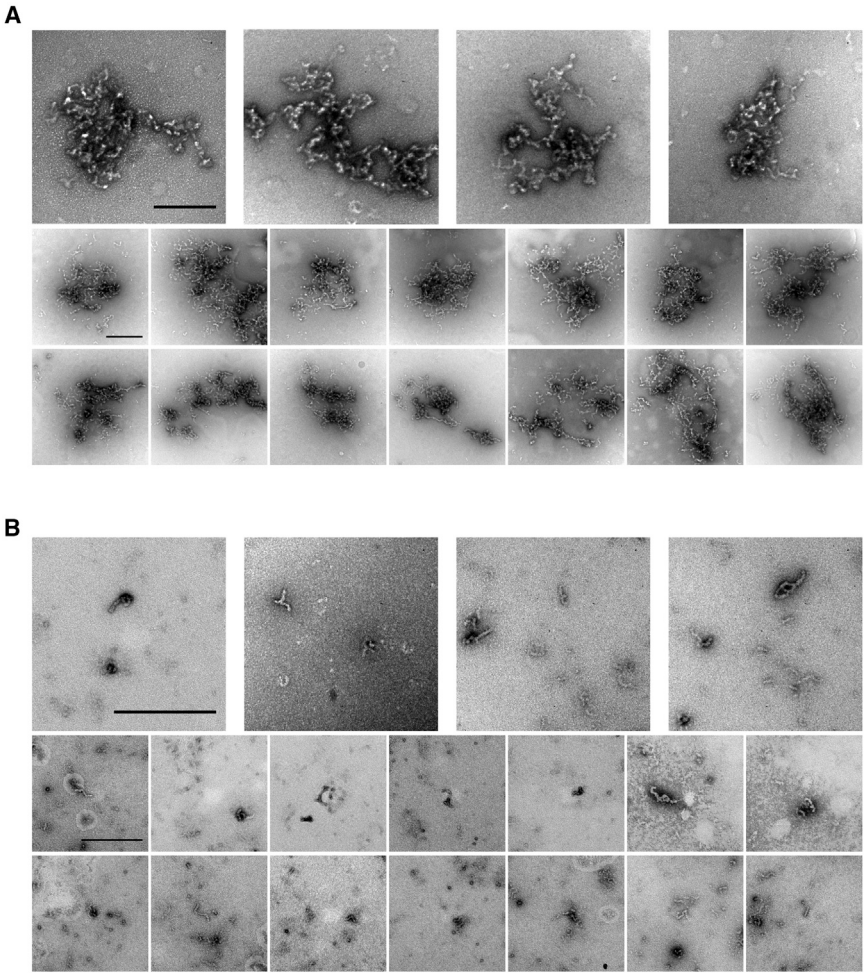

Figure 6. DNA-Dependent Enhancement of Fiber Formation of FUS-FLI Fusion Protein

(A) mCherry:FUS LC domain-FLI DNA-binding domain fusion protein (0.5 μM) was incubated with microsatellite DNA (20 nM) in the presence of 35% glycerol. After 1 hr incubation, materials in the solution were visualized by transmission electron microscopy. FUS-FLI fibers grew long and became large sleave.

(B) mCherry:FUS-FLI protein was incubated in the presence of 35% glycerol (no DNA). Small amounts of spontaneous nucleation and fiber growth were observed. All scale bars indicate 0.5 μm.

See also Figure S5.

to different DNA-binding domains. Beautiful ChIP-chip experiments have given evidence that when the LC domain of EWS is linked to the FLI DNA-binding domain, it binds in living cells to microsatellites carrying iterative repeats of the tetranucleotide sequence GGAA (Gangwal et al., 2008). Since GGAA represents an optimal binding site for the ETS DNA-binding domain of the EWS:FLI protein, the fusion protein binds iteratively to these otherwise inert microsatellites. If located within 5–10 kb of a target promoter, the translocation product-bound microsatellite has been hypothesized to become a rogue enhancer. We are like-minded with the Cech lab in providing evidence that enhanced local concentration via proximal binding of multiple subunits of LC domain containing proteins will represent a critical stimulus for polymerization (Figure 6).

Despite coming to similar conclusions regarding LC polymerization and CTD binding, the Cech studies differ in many ways from those reported herein. Cech and colleagues have employed intact FUS protein and have studied the protein in the context of its normal cellular function. We, instead, have focused on the isolated parts of FET proteins that are translocated in cancer onto a variety of DNA-binding domains. It is relatively straightforward to conceptualize how the latter class of proteins bind DNA, polymerize locally via their LC domains, and thereby achieve the fibrous organizational state required to bind the CTD of RNA polymerase II. How, instead, might one conceptualize the manner in which native, RNA-binding forms of FET proteins achieve the same task? One possibility is that native FET proteins are able to bind to noncoding RNAs that remain nascently attached to DNA. Perhaps these complexes might function as "RNA enhancers," (Kim et al., 2010; Lam et al., 2013; Li et al., 2013; Wang et al., 2011) achieving the goal of bringing FET proteins into the proximity of target genes at locally elevated protein concentrations adequate to facilitate LC domain polymerization necessary for the recruitment of RNA polymerase II.

Even if both of these conceptualizations of polymer-dependent CTD recruitment are fundamentally correct, we are undoubtedly only beginning to see the tip of the iceberg. Hundreds of RNA regulatory proteins and transcription factors contain LC domains, offering almost limitless opportunity for complexity in cellular organization—including the possible existence of solid-state polymeric pathways for information flow from "transcription factories" through fibrous nuclear bodies into various forms of cytoplasmic RNA granules to the ultimate sites of mRNA translation. Anticipating that LC domains are almost certainly regulated by posttranslational modification, it is clear that we are at the earliest stage of recognizing this new and unconventional way of thinking about cellular organization. Indeed, if we are correct in assuming that phosphorylation of the CTD of RNA polymerase II triggers its release from FET protein fibers, this may serve as an initial paradigm for considering posttranslational modification as a means of controlling the dynamic behavior of LC derived polymers.

We close with three perplexing observations. First, why do polymeric fibers formed from the LC domain of TAF15 bind the CTD so much more avidly than the analogous LC domains of FUS and EWS (Figure 3A)? Second, why is it that mutations of

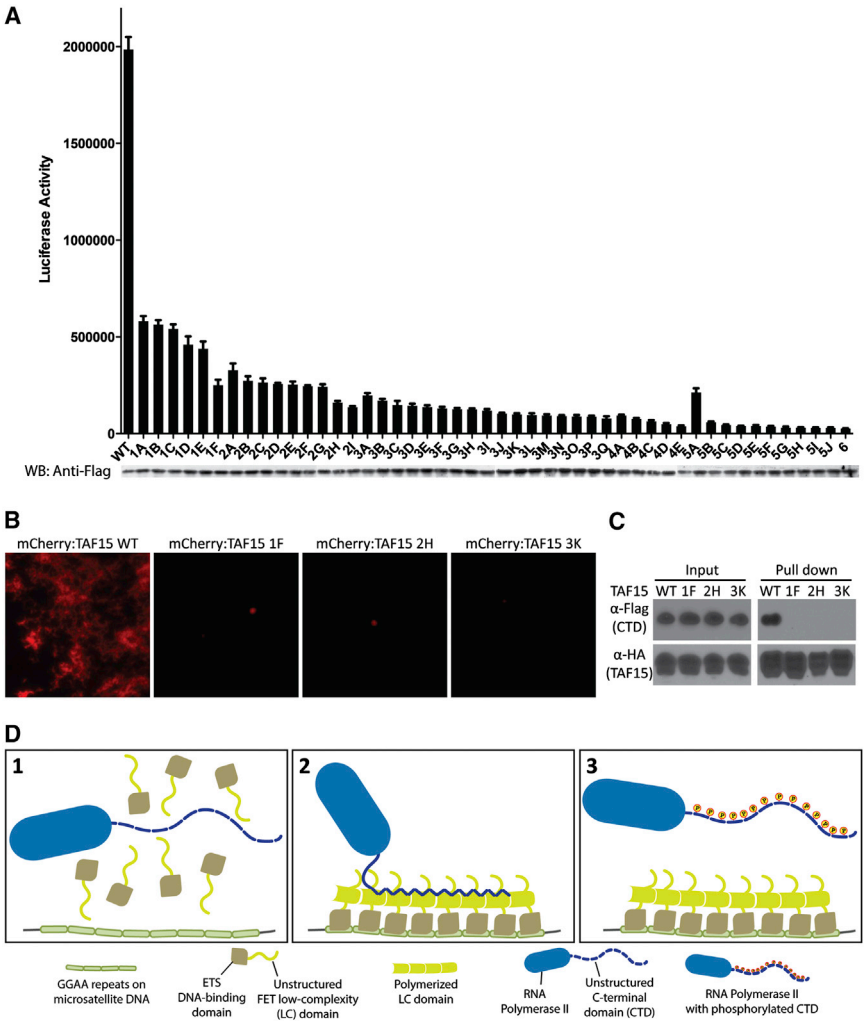

Figure 7. Polymerization of the LC Domain of TAF15 is Required for Both Transcriptional Activation and CTD Binding

(A) Forty-eight tyrosine-to-serine mutations were randomly introduced into the LC domain of TAF15 (Table S4). Mutants were assayed for transcriptional activation capacity as GAL4 fusion proteins. Expression levels for individual test proteins were monitored by western blotting as displayed below histogram depictions of transcriptional activation measurements.

(B) mCherry fusion proteins linked to the native LC domain of TAF15 (WT), the 1F mutant, the 2H mutant, and the 3K mutant were incubated under conditions favorable for polymerization (Experimental Procedures). Fluorescence microscopy was employed as an assay for fiberization of the four test proteins. See also Movie S2.

(C) Coimmunoprecipitation assays were conducted by mixing Flag-tagged GFP:CTD_{C26} with HA-tagged versions of mCherry linked to the native form of the TAF15 LC domain, or to the 1F, 2H or 3K tyrosine-to-serine mutants. Following HA-mediated immunoprecipitation, samples were run on a denaturing SDS-PAGE gel and subjected to western blotting using either anti-Flag or anti-HA antibodies. See also Figure S6.

(D) Schematic Concept of RNA Polymerase II Recruitment by LC Domain Polymers. (1) The CTD of RNA polymerase II does not bind to monomers of the unstructured LC domain of FET protein fused to the ETS DNA-binding domain. (2) Once the ETS DNA-binding domain of the fusion protein binds to the GGAA repeats on microsatellite DNA, the LC domains of FET protein form fibrous polymer that can recruit RNA polymerase II via direct interaction with the CTD. (3) Phosphorylation of serine residues 2, 5, and 7 of the CTD heptad repeats by CDK7 or CDK9 facilitates release of RNA polymerase II from the FET LC domain polymer.

two tyrosines in the LC domain of FUS yield a protein with enhanced transcriptional activation capacity and an enhanced ability to bind the CTD of RNA polymerase II (Figures 1 and 5)? Third, why do the degenerate CTD repeats bind the polymeric fibers of FET LC domains better than the consensus repeats (Figure 3B)? Since TAF15 was discovered as a component of the TFIID complex, one might imagine that its function has become more proximally evolved to the transcription initiation process than FUS or EWS, thus consistent with constitutive, avid affinity for the CTD of RNA polymerase II. The region of TAF15's LC domain most distinctly different from the LC domains of FUS and EWS is in the [G/S]Y[G/S] repeats 16-21, precisely the same region of the location of the two tyrosine-to-serine mutations in the 2A gain-of-function mutant of FUS (Figure 5). Six straight repeats of TAF15 contain aspartic acid on one side or the other of the central tyrosine. We speculate that these aspartic acid residues disrupt or in some way modify the TAF15 polymeric fiber in this region, and that such modifications enhance CTD binding. Likewise, we hypothesize that the two tyrosine-to-serine alterations in the 2A mutant of the FUS LC domain also modify the polymeric fiber in a manner favoring CTD binding and transcriptional activation capacity. Perhaps FUS and EWS can be modified by phosphorylation on serine residues flanking the tyrosines in this "16-21 window" of their LC domains, such that fiber disruption and propensity for CTD binding might be controlled in a regulated manner—rather than being "hard-wired" as is possible owing to TAF15's evolutionary accumulation of aspartic acid residues as phosphomimetics.

Finally, we consider why it might be that the polymeric form of the TAF15 LC domain binds the C-terminal degenerate CTD repeats so much more strongly than the N-terminal consensus CTD repeats (Figure 3B). The N-terminal consensus repeats of the CTD have been shown to bind and cocrystallize with components of the mediator complex (Robinson et al., 2012). Likewise, elegant EM studies have also provided evidence of CTD:mediator interaction (Tsai et al., 2013). Perhaps upon evolutionary duplication of the CTD from 26 repeats in yeast to 52 repeats in mammals has allowed the degenerate repeats to adopt a new function molded to the fibrous LC domains of FET proteins. Close inspection of the degenerate repeats may be revealing in this regard. In comparing the degenerate repeats to our most distant evolutionary relative bearing the degenerate repeats,

teleost fish from which humans diverged 450 million years ago, it is notable that the pattern of degeneracy is nearly identical. Briefly put, the degenerate changes installed nearly one half billion years ago have been perfectly conserved. The most notable difference between the degenerate and consensus CTD repeats is the presence of lysine residues replacing S_7 ($Y_1S_2P_3T_4S_5P_6\mathbf{S_7}$ versus $Y_1S_2P_3T_4S_5P_6\mathbf{K_7}$) in seven of the degenerate repeats. Although embarrassingly simplistic, we speculate that the positive charge of these seven lysine residues on the immediate N-terminal sides of tyrosines in the degenerate CTD repeats might blend with the negative charge of six consecutive aspartic acids flanking tyrosines in the "16-21 window" of TAF15. If correct, this idea of charge attraction might explain the longstanding enigma of acidic activation domains.

EXPERIMENTAL PROCEDURES

Materials

Synthesis of the biotinylated isoxazole is described previously (Kato et al., 2012). Mammalian pM expression vector was purchased from Clontech Laboratory (USA). CDK7/Cyclin H/MAT1 and CDK9/Cyclin T1 enzyme complexes were purchased from Millipore.

Transcriptional Activation by Tyrosine-to-Serine Mutants of FUS and TAF15 LC Domain

Mammalian expression plasmids for human FUS or TAF15 LC domain fused to GAL4 DNA-binding domain were constructed by insertion of a DNA fragment encoding FUS LC domain (residues 2-266) or TAF LC domain (residues 2-208) into the multiple cloning sites of pM vector (Clontech, USA). Tyrosine-to-serine mutants of GAL4-FUS LC and GAL4-TAF15 LC were generated by QuickChange Multi Site-Directed Mutagenesis Kit (Agilent Technologies). To test transcriptional activity of GAL4-fused LC domains, luciferase assays were performed in quadruplicate in 96-well plate. In each well, 50 ng of pM-LC vector and 20 ng L8G5-Luc (Lemercier et al., 2000) were reverse transfected to HEK293T cells (6,500 cells per well) by lipofectamine 2000 (Invitrogen). Luciferase assays were performed after 24 hr of incubation at 37°C, using Bright-Glo Luciferase Assay System (Promega). In order to obtain similar protein levels relative to wild-type, 25 ng, 50 ng, or 100 ng of pM-LC mutant plasmids were cotransfected with L8G5-Luc, and expression of GAL4-fused LC domains was determined by western blotting with anti-GAL4 antibody (sc-510, Santa Cruz Biotechnology). The DNA dosages yielding the same expression level of GAL4-FUS mutant proteins to the wild-type protein were used for luciferase assays.

Generation of Bacterial Expression Constructs

Expression plasmids for recombinant proteins used in all the biochemical assays were constructed using pHis-parallel vectors (Sheffield et al., 1999) as described before (Kato et al., 2012). The details were described in Supplemental Information.

Protein Purification

All recombinant proteins were overexpressed in *E. coli* strain BL21(DE3). mCherry and GFP fusion LC domains from FUS, EWS, TAF15, CIRBP and hnRNPA2 were expressed and purified as described previously (Kato et al., 2012). GFP:CTD proteins were purified by Ni-NTA (QIAGEN, USA) in the same conditions described previously (Kato et al., 2012), and further purified by a Hi-Load Superdex 200 gel filtration column (GF Healthcare, USA). The expression and purification of his-tagged mCherry:FUS LC domain-FLI DNA-binding domain fusion protein is described in Supplemental Information.

Hydrogel-Binding Assays

Hydrogel droplets of mCherry:LC domain of either FUS, EWS, TAF15, CIRBP, or hnRNPA2 were prepared as described before (Kato et al., 2012). Briefly, concentrated mCherry fusion proteins (typically ~50 mg/ml) were dialyzed in gelation buffer containing 20 mM Tris-HCl (pH 7.5), 200 mM NaCl, 20 mM BME, 0.5 mM EDTA, 0.1 mM PMSF over night. Dialyzed protein solution was sonicated, clarified by centrifugation, and concentrated again to ~50–80 mg/ml. Small droplets (0.5 μL) of this protein solution were placed on a glass-bottomed dish (MatTek, USA). Hydrogels typically form after incubation for a couple of days at room temperature. For standard hydrogel-binding assays, glycerol stocks of the purified GFP fusion test proteins were diluted to 1 μM in 1 ml of the gelation buffer. The GFP test solution was pipetted into the hydrogel dish so as to soak the hydrogel droplets in the GFP solution. After overnight incubation at 4°C, a horizontal section of the hydrogel droplet was scanned with both the mCherry and GFP excitation wavelengths on a Zeiss LSM510 confocal microscope. To measure binding rates of the FUS LC domain mutants carrying different number of tyrosine-to-serine mutations onto the wild-type mCherry:FUS hydrogels, immediately after GFP:mutant solution was administered into the hydrogel dish, scans of a horizontal section of the hydrogel droplet were performed at indicated time points. GFP signals of the scanned sections were integrated by the program ImageJ (Rasband, 1997-2011). Initial binding rates were calculated by a linear regression fitting with the program Prism (Graphpad, USA). The scatter plot for calculating correlation between transactivation and hydrogel binding rate of the mutants was made using the program Prism.

Cyclin-Dependent Kinase Assay

For release of GFP:CTD_{C26} from mCherry:TAF15, mCherry:FUS and mCherry:FUS 2A hydrogels by phosphorylation, unphosphorylated GFP:CTD_{C26} was preincubated with the hydrogels overnight. After removing the GFP:CTD_{C26} solution, the CDK7 or CDK9 reaction mixtures were applied to the GFP:CTD_{C26}-trapped hydrogel droplets. The hydrogel plates were incubated at 30°C for 1 hr. Hydrogels were analyzed by a confocal microscope. For western blotting of the released protein from the hydrogel droplets, the reaction mixture was carefully removed from the hydrogel plate and mixed with 2 × SDS lysis buffer. After SDS-PAGE, GFP:CTD_{C26} protein was transferred to a nitrocellulose filter and visualized by western blotting using anti-GFP antibodies (Chemicon) or the 3E8 phospho-S_5 antibody (Millipore).

Biotinylated Isoxazle-Mediated Precipitation

Protein precipitation by b-isox chemical from yeast extract and human nuclear extract is described in Supplemental Information.

Cloning of Microsatellite DNA

The microsatellite DNA in the promoter region of the hNROB1 gene was amplified by PCR from U2OS cell genomic DNA as a template. The sequence of the microsatellite DNA is shown in Supplemental Information.

Transmission Electron Microscopy

DNA-dependent fiber formation of mCherry:FUS-FLI protein was carried out in a reaction mixture containing 20 mM Tris-HCl (pH 7.5), 200 mM NaCl, 20 mM BME, 0.5 mM EDTA, 0.1 mM PMSF, and 35% glycerol, with or without 20 nM microsatellite DNA. mCherry:FUS-FLI protein stored in 50% glycerol was added to the reaction mixture at the final concentration of 0.5 μM, and the reactions were incubated at room temperature for 1 hr. The sample solution (10 μL) was spotted on a TEM grid (FCF-200-Cu Copper grid from Electron Microscopy Sciences, USA). After blotting the excess solution, the surface of the grid was washed with 10 μl distilled water twice. After removing the excess water, the grid was stained for 10 s with 10 μl drop of 2% uranyl acetate in water. After the uranyl solution was blotted, the grid was dried in air. TEM images were obtained at 120 kV on Tecnai TEM.

Fiber Extension Assays

The mCherry:TAF15 WT, 1F, 2H, or 3K protein solution (100 μM) in the gelation buffer was sonicated to break down the preformed fibers, if any, into short fiber seeds. After spin down at 14K rpm for 5 min at 4°C, the protein solution was transferred to new test tubes. After overnight incubation at 4°C, the proteins were 10-fold diluted in the gelation buffer and the formed fibers were detected by fluorescence microscopy.

Pull Down of GFP:CTD$_{C26}$ by mCherry:TAF15

HA- or Flag-tag was added, respectively, to the C terminus of mCherry:TAF15 or GFP:CTD$_{C26}$ by PCR amplification. mCherry:TAF15 WT, 1F, 2H, or 3K-HA proteins were diluted to 10 μM in gelation buffer with 0.05% NP-40. After 4 hr of incubation at 4°C, 1 μM of GFP:CTD$_{C26}$-Flag and 10 μl of HA-magnetic beads (Pierce, USA) were added to each tube containing mCherry:TAF15 solution. The mixtures were incubated for overnight at 4°C with gentle rotation. The magnetic beads were washed twice with the gelation buffer with 0.05% NP-40. Protein was recovered by boiling at 95°C for 10 min in 2× SDS lysis buffer. Protein was visualized by western blotting using Flag or HA antibodies.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, six figures, four tables, and two movies and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2013.10.033.

ACKNOWLEDGMENTS

We thank the shared resource cores in Microscopic Imaging, Proteomics and DNA Sequencing at UTSWMC for technical support; Dr. Robert Tycko for provision of the electron micrograph of amyloid fibers included in the graphical abstract; Dr. David Trudgian for help with mass spectrometry; and Drs. Bruce Alberts, Deepak Nijhawan, Michael Brown, and Michael Rosen for helpful comments on the composition of the manuscript. This work was funded by an unrestricted endowment provided S.L.M. by an anonymous donor and NIH grant GM066108 for J.L.C.

Received: August 1, 2013
Revised: September 13, 2013
Accepted: October 1, 2013
Published: November 21, 2013

REFERENCES

Akhtar, M.S., Heidemann, M., Tietjen, J.R., Zhang, D.W., Chapman, R.D., Eick, D., and Ansari, A.Z. (2009). TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. Mol. Cell *34*, 387–393.

Andersson, M.K., Ståhlberg, A., Arvidsson, Y., Olofsson, A., Semb, H., Stenman, G., Nilsson, O., and Aman, P. (2008). The multifunctional FUS, EWS and TAF15 proto-oncoproteins show cell type-specific expression patterns and involvement in cell spreading and stress response. BMC Cell Biol. *9*, 37.

Arvand, A., and Denny, C.T. (2001). Biology of EWS/ETS fusions in Ewing's family tumors. Oncogene *20*, 5747–5754.

Bertolotti, A., Bell, B., and Tora, L. (1999). The N-terminal domain of human TAFII68 displays transactivation and oncogenic properties. Oncogene *18*, 8000–8010.

Buratowski, S. (2009). Progression through the RNA polymerase II CTD cycle. Mol. Cell *36*, 541–546.

Chapman, R.D., Heidemann, M., Albert, T.K., Mailhammer, R., Flatley, A., Meisterernst, M., Kremmer, E., and Eick, D. (2007). Transcribing RNA polymerase II is phosphorylated at CTD residue serine-7. Science *318*, 1780–1782.

Corden, J.L., Cadena, D.L., Ahearn, J.M., Jr., and Dahmus, M.E. (1985). A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. Proc. Natl. Acad. Sci. USA *82*, 7934–7938.

Crozat, A., Aman, P., Mandahl, N., and Ron, D. (1993). Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma. Nature *363*, 640–644.

Cumberworth, A., Lamour, G., Babu, M.M., and Gsponer, J. (2013). Promiscuity as a functional trait: intrinsically disordered regions as central players of interactomes. Biochem. J. *454*, 361–369.

Czudnochowski, N., Bösken, C.A., and Geyer, M. (2012). Serine-7 but not serine-5 phosphorylation primes RNA polymerase II CTD for P-TEFb recognition. Nat Commun *3*, 842.

Egloff, S., and Murphy, S. (2008). Cracking the RNA polymerase II CTD code. Trends Genet. *24*, 280–288.

Egloff, S., Dienstbier, M., and Murphy, S. (2012). Updating the RNA polymerase CTD code: adding gene-specific layers. Trends Genet. *28*, 333–341.

Gangwal, K., Sankar, S., Hollenhorst, P.C., Kinsey, M., Haroldsen, S.C., Shah, A.A., Boucher, K.M., Watkins, W.S., Jorde, L.B., Graves, B.J., and Lessnick, S.L. (2008). Microsatellites as EWS/FLI response elements in Ewing's sarcoma. Proc. Natl. Acad. Sci. USA *105*, 10149–10154.

Guipaud, O., Guillonneau, F., Labas, V., Praseuth, D., Rossier, J., Lopez, B., and Bertrand, P. (2006). An in vitro enzymatic assay coupled to proteomics analysis reveals a new DNA processing activity for Ewing sarcoma and TAF(II)68 proteins. Proteomics *6*, 5962–5972.

Han, T.W., Kato, M., Xie, S., Wu, L.C., Mirzaei, H., Pei, J., Chen, M., Xie, Y., Allen, J., Xiao, G., and McKnight, S.L. (2012). Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. Cell *149*, 768–779.

Huntley, M.A., and Golding, G.B. (2002). Simple sequences are rare in the Protein Data Bank. Proteins *48*, 134–140.

Ichikawa, H., Shimizu, K., Katsu, R., and Ohki, M. (1999). Dual transforming activities of the FUS (TLS)-ERG leukemia fusion protein conferred by two N-terminal domains of FUS (TLS). Mol. Cell. Biol. *19*, 7639–7650.

Jones, J.C., Phatnani, H.P., Haystead, T.A., MacDonald, J.A., Alam, S.M., and Greenleaf, A.L. (2004). C-terminal repeat domain kinase I phosphorylates Ser2 and Ser5 of RNA polymerase II C-terminal domain repeats. J. Biol. Chem. *279*, 24957–24964.

Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J., et al. (2012). Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. Cell *149*, 753–767.

Kim, T.K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin, D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., et al. (2010). Widespread transcription at neuronal activity-regulated enhancers. Nature *465*, 182–187.

Kim, M., Suh, H., Cho, E.J., and Buratowski, S. (2009). Phosphorylation of the yeast Rpb1 C-terminal domain at serines 2, 5, and 7. J. Biol. Chem. *284*, 26421–26426.

Lam, M.T., Cho, H., Lesch, H.P., Gosselin, D., Heinz, S., Tanaka-Oishi, Y., Benner, C., Kaikkonen, M.U., Kim, A.S., Kosaka, M., et al. (2013). Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. Nature *498*, 511–515.

Lemercier, C., Verdel, A., Galloo, B., Curtet, S., Brocard, M.P., and Khochbin, S. (2000). mHDA1/HDAC5 histone deacetylase interacts with and represses MEF2A transcriptional activity. J. Biol. Chem. *275*, 15594–15599.

Lessnick, S.L., and Ladanyi, M. (2012). Molecular pathogenesis of Ewing sarcoma: new therapeutic and transcriptional targets. Annu. Rev. Pathol. *7*, 145–159.

Li, W., Notani, D., Ma, Q., Tanasa, B., Nunez, E., Chen, A.Y., Merkurjev, D., Zhang, J., Ohgi, K., Song, X., et al. (2013). Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature *498*, 516–520.

Phatnani, H.P., and Greenleaf, A.L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. Genes Dev. *20*, 2922–2936.

Rabbitts, T.H., Forster, A., Larson, R., and Nathan, P. (1993). Fusion of the dominant negative transcription regulator CHOP with a novel gene FUS by translocation t(12;16) in malignant liposarcoma. Nat. Genet. *4*, 175–180.

Ramanathan, Y., Rajpara, S.M., Reza, S.M., Lees, E., Shuman, S., Mathews, M.B., and Pe'ery, T. (2001). Three RNA polymerase II carboxyl-terminal domain kinases display distinct substrate preferences. J. Biol. Chem. *276*, 10913–10920.

Rasband, W.S. (1997-2011). ImageJ (Bethesda: U. S. National Institute of Health).

Rickert, P., Corden, J.L., and Lees, E. (1999). Cyclin C/CDK8 and cyclin H/CDK7/p36 are biochemically distinct CTD kinases. Oncogene *18*, 1093–1102.

Robinson, P.J., Bushnell, D.A., Trnka, M.J., Burlingame, A.L., and Kornberg, R.D. (2012). Structure of the mediator head module bound to the carboxy-terminal domain of RNA polymerase II. Proc. Natl. Acad. Sci. USA *109*, 17931–17935.

Roy, R., Adamczewski, J.P., Seroz, T., Vermeulen, W., Tassan, J.P., Schaeffer, L., Nigg, E.A., Hoeijmakers, J.H., and Egly, J.M. (1994). The MO15 cell cycle kinase is associated with the TFIIH transcription-DNA repair factor. Cell *79*, 1093–1101.

Schwartz, J.C., Ebmeier, C.C., Podell, E.R., Heimiller, J., Taatjes, D.J., and Cech, T.R. (2012). FUS binds the CTD of RNA polymerase II and regulates its phosphorylation at Ser2. Genes Dev. *26*, 2690–2695.

Schwartz, J.C., Wang, X., Podell, E.R., and Cech, T.R. (2013). RNA seeds higher-order assembly of FUS protein. Cell Rep. *5*. Published online November 21, 2013. http://dx.doi.org/10.1016/j.celrep.2013.11.017.

Sheffield, P., Garrard, S., and Derewenda, Z. (1999). Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. Protein Expr. Purif. *15*, 34–39.

Trigon, S., Serizawa, H., Conaway, J.W., Conaway, R.C., Jackson, S.P., and Morange, M. (1998). Characterization of the residues phosphorylated in vitro by different C-terminal domain kinases. J. Biol. Chem. *273*, 6769–6775.

Tsai, K.L., Sato, S., Tomomori-Sato, C., Conaway, R.C., Conaway, J.W., and Asturias, F.J. (2013). A conserved Mediator-CDK8 kinase module association regulates Mediator-RNA polymerase II interaction. Nat. Struct. Mol. Biol. *20*, 611–619.

Uversky, V.N. (2002). Natively unfolded proteins: a point where biology waits for physics. Protein Sci. *11*, 739–756.

Wang, D., Garcia-Bassets, I., Benner, C., Li, W., Su, X., Zhou, Y., Qiu, J., Liu, W., Kaikkonen, M.U., Ohgi, K.A., et al. (2011). Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature *474*, 390–394.

Zinszner, H., Albalat, R., and Ron, D. (1994). A novel effector domain from the RNA-binding protein TLS or EWS is required for oncogenic transformation by CHOP. Genes Dev. *8*, 2513–2526.

EXTENDED EXPERIMENTAL PROCEDURES

Construction of Bacterial Expression Plasmids

Construction of expression plasmids for mCherry or GFP fusion LC domains of FUS (residues 2-214), CIRBP (full length) and hnRNPA2 (residues 181-341) was described previously (Kato et al., 2012). Expression plasmids for LC domains of EWS and TAF15 were constructed using similar methods as described (Kato et al., 2012). DNA fragments encoding EWS LC domain (residues 47-266) or TAF15 LC domain (residues 2-208) were amplified by PCR using a cDNA library made from human U2OS cells as a template and subcloned into the pHis-mCherry parallel vector. To generate bacterial expression constructs of RNA-dependent polymerase II C-terminal domain (Poll II CTD), DNA fragments encoding different size of Pol II CTD (C26, heptapeptide repeats 27-52; C20, 33-52; C15, 38-52; C10, 43-52; C5, 48-52; N20, 1-20) were also obtained by PCR amplification from a U2OS cDNA library. The DNA fragments of the Pol II CTD were inserted into the pHis-GFP parallel vector. To construct the mCherry:FUS-FLI fusion gene plasmid, DNA fragments encoding FUS LC domain (residues 2-214) and the FLI C-terminal DNA-binding domain (residues 220-452) were fused by a bridging PCR technique. The fused DNA fragment was inserted into pHis-mCherry parallel vector.

Protein Purification

The expression plasmid for his-tagged mCherry:FUS LC domain-FLI DNA-binding domain fusion protein was transformed into *E. coli* BL21 (DE3). Cells were inoculated in 50 ml LB media from a single colony and cultured over-night at 30°C without shaking. The following day the 10 ml preculture media was transferred into 1 L M9 media and cultured at 37°C for 8 hr. Protein expression was then induced by administration of 0.5 mM IPTG, and the cells were cultured at 16°C overnight. Cells were harvested and stored at −80°C. Protein was purified under denaturing conditions. First, the cells were lysed in a urea lysis buffer containing 8 M urea, 100 mM NaH_2PO_4 and 10 mM Tris, (pH 8.0) over-night. The cell lysate was sonicated then centrifuged at 35,000 rpm at 4°C for 1 hr. Supernatant was mixed with Ni-NTA resin (QIAGEN, USA) for 30 min at room temperature. The Ni-NTA resin was packed in a glass column and washed with urea wash buffer containing 8 M urea, 100 mM NaH_2PO_4 and 10 mM Tris (pH 6.3). Bound proteins were eluted from the resin with urea elution buffer containing 8 M urea, 100 mM NaH_2PO_4 and 10 mM Tris (pH 4.5). To refold eluted proteins, the elution was diluted 100-fold with heparin buffer A containing 20 mM Tris-HCl (pH 7.5), 20 mM β-mercaptoethanol (BME), and 0.1 mM PMSF. After removing insoluble proteins by centrifugation at 20,000 rpm for 20 min at 4°C, the supernatant was loaded on a heparin sepharose column (GE Healthcare, USA). The column was washed with heparin wash buffer containing 20 mM Tris-HCl (pH 7.5), 20 mM BME, 0.1 mM PMSF, and 100 mM NaCl. Bound proteins were eluted with heparin elution buffer containing 20 mM Tris-HCl (pH 7.5), 20 mM BME, 0.5 mM EDTA, 0.1 mM PMSF, and 500 mM NaCl. Glycerol was added to the elution at a final concentration of 50%. The protein solution was kept at −20°C. The purities of the purified proteins were confirmed by SDS-PAGE, and the concentrations were determined by absorbance at UV_{280}. mCherry fusion proteins were stored at −80°C after purification and concentration. GFP fusion proteins were stored at −20°C as 50% glycerol stock after purification.

Sequence of Microsatellite DNA

As shown below, the microsatellite DNA (294 bp) contains 25 repeats of GGAA fragment (underlined) flanked by 100 nonrepeated nucleotides. TAGGTTTTCCTCTTATGCTGAGAATTCCAGGTCCTGGAGAAGAAGAAAAAGAGAAAGAAAGAGAGAGAGAGAAGGAG
TGAGAGAGGGAGGGAGGGAGGGAGGGAGGGAGGAAGGAAGGAAGGAAGGAAGGAAGGAAAGGAAGGAAGGAAGGAAGGAAG
GAAGGAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAAAGAAACAGCAAAAAAAGAAAGAGGGAGGAT
GGGAGGGAGGGAAAAAGTAAAAATGATTCTGTATCAGCTGGTATATACCAACA

Biotinylated Isoxazle-Mediated Precipitation of Yeast Extracts

Wild-type *S. cerevisiae* were grown in YPD to 1 OD_{600}. Cells were spun down and resuspended in lysis buffer containing 200 mM Tris-HCl (pH 8), 320 mM ammonium sulfate, 5 mM $MgCl_2$, 10 mM EGTA, 20 mM EDTA, 1 mM dithiothreitol, 20% glycerol, plus protease inhibitors (RPI) (Patturajan et al., 1998). After breaking cells in a FastPrep (MP Biomedicals) with glass beads, debris were removed by centrifugation and the supernatant was desalted into isox buffer containing 20 mM Tris, 150 mM NaCl, 5 mM $MgCl_2$, 20 mM BME, 0.5% NP-40, 10% glycerol + protease inhibitors (Kato et al., 2012) using a spin column (Thermo). B-Isox precipitation was for 1 hr at 5°C with 3 μl of 1 mM b-isox/100μl of extract. Pellets were washed three times with the isox buffer, resuspended in SDS sample buffer, separated on a 4%–12% tris-glycine gel, blotted and probed with anti-CTD antibodies (Chapman et al., 2007; Thompson et al., 1990).

Biotinylated Isoxazole-Mediated Precipitation of Mammalian Nuclear Extracts

To prepare nuclear extract, ten 15-cm plates of HEK293FT cells were collected by trypsinization, washed in 1× PBS, and resuspended in hypotonic buffer A (50 mM Tris-HCl (pH 7.9), 3 mM $MgCl_2$, 10 mM NaCl, 0.2 mM PMSF, 1:300 P8340, Sigma, 20 mM BME). The cells were incubated on ice for 60 min, homogenized with a 40 ml glass dounce and then spun 1,000 × *g* for 15 min. The pellet was then resuspended in hypotonic buffer A and sonicated using a bioruptor for 5 min twice (30 s on/30 s off ice). The sonicated resuspension was then spun at 15,000 × *g* for 15 min. The supernatant was transferred to a fresh tube, NaCl was added to a final concentration of 150 mM, and 0.5% NP-40 was added. The supernatant was then incubated with 100 μM b-isox, allowed to rotate at 4C° for 2 hr, and then centrifuged 15,000 × *g* for 15 min to pellet the b-isox precipitant. The pellet washed once in buffer A,

resuspended in 2× SDS loading buffer for SDS-PAGE analysis. The entire lane was excised from the gel and submitted for identification by mass spectrometry analysis.

Biotinylated Isoxazle-Mediated Precipitation of HeLa Cells

HeLa cells were grown in YPD to 1 OD_{600}. Cells were spun down and resuspended in lysis buffer containing 200 mM Tris-HCl (pH 8), 320 mM ammonium sulfate, 5 mM $MgCl_2$, 10 mM EGTA, 20 mM EDTA, 1 mM dithiothreitol, 20% glycerol, plus protease inhibitors (RPI) (Patturajan et al., 1998). After breaking cells in a FastPrep (MP Biomedicals) with glass beads, debris were removed by centrifugation and the supernatant was desalted into B-isox buffer containing 20 mM Tris, 150 mM NaCl, 5 mM $MgCl_2$, 20 mM BME, 0.5% NP-40, 10% glycerol + protease inhibitors (Kato et al., 2012) using a spin column (Thermo). B-Isox precipitation was for 1 hr at 5°C with 3 μl of 1 mM isox/ 100 μl of extract. Pellets were washed three times with the isox buffer, resuspended in SDS sample buffer, separated on a 4%–12% Tris-glycine gel, blotted and probed with anti-CTD antibodies (Chapman et al., 2007; Thompson et al., 1990).

Biotinylated Isoxazole-Mediated Precipitation of GFP:CTD_{C26}

GFP:CTD_{C26} was phosphorylated by either CDK7 or CDK9. Protein was diluted to 1 μM in B-isox buffer and then biotinylated Isoxazole was added at 100 μM. After 1 hr incubation with gentle rotation at 4°C, the reaction was spun down at 14,000 rpm for 15 min to pellet the precipitates. The pellet was washed twice with the reaction buffer and resuspended in 2× SDS sample loading buffer. After SDS-PAGE, protein was detected by western blotting using an anti-GFP antibody (Chemicon, USA).

CTD Binding to Polymeric Fibers of mCherry:TAF15

mCherry:TAF15 polymeric fibers were prepared as described before (Kato et al., 2012). The fiber solution was diluted in gelation buffer (1 μM final monomeric concentration). GFP-CTD_{C26} or GFP alone was added to this solution (2.6 μM final concentration). The mixture was immediately centrifuged at 18,000 rpm for 1 min to precipitate fibers. After the supernatant was carefully removed, fibers were resupended in the fresh gelation buffer. The washing step was repeated one more time. Washed fibers were subsequently visualized by fluorescence microscopy.

Cyclin-Dependent Kinase Assay

For phosphorylation of GFP:CTD_{C26}, 10 μM of purified protein was incubated with indicated amounts of CDK7 or CDK9 in CDK reaction buffer containing 50 mM Mops-NaOH (pH 7.0), 1 mM EDTA, 0.001% NP-40, 2.5% glycerol, 0.05% BME and 10 mM MgAc in the presence of 0.4 μM ATP at room temperature for 2 hr. For mCherry:TAF15 hydrogel binding, reaction mixtures were 10-fold diluted into the gelation buffer to make 1 μM of final concentration of GFP:CTD_{C26}. Diluted reaction mixtures were then applied to the mCherry:TAF15 hydrogel droplets and incubated at 4°C for overnight. Hydrogels were analyzed by a confocal microscope.

DNA Gel Shift Assays

Microsatellite DNA (294 bp, 0.1 μM) was mixed with indicated concentrations of mCherry:FUS LC domain-FLI DNA-binding domain fusion protein in a buffer containing 10 mM Tris-HCl (pH 7.5), 100 mM NaCl, 10 mM BME, 0.25 mM EDTA, 0.05 mM PMSF, and 40% glycerol, and then incubated for 1 hr on ice. The reaction mixtures were separated by 1.3% agarose gel followed by ethidium bromide staining. The stained gel was visualized by a UV transilluminator.

SUPPLEMENTAL REFERENCES

Patturajan, M., Schulte, R.J., Sefton, B.M., Berezney, R., Vincent, M., Bensaude, O., Warren, S.L., and Corden, J.L. (1998). Growth-related changes in phosphorylation of yeast RNA polymerase II. J. Biol. Chem. *273*, 4689–4694.

Thompson, N.E., Aronson, D.B., and Burgess, R.R. (1990). Purification of eukaryotic RNA polymerase II by immunoaffinity chromatography. Elution of active enzyme with protein stabilizing agents from a polyol-responsive monoclonal antibody. J. Biol. Chem. *265*, 7069–7077.

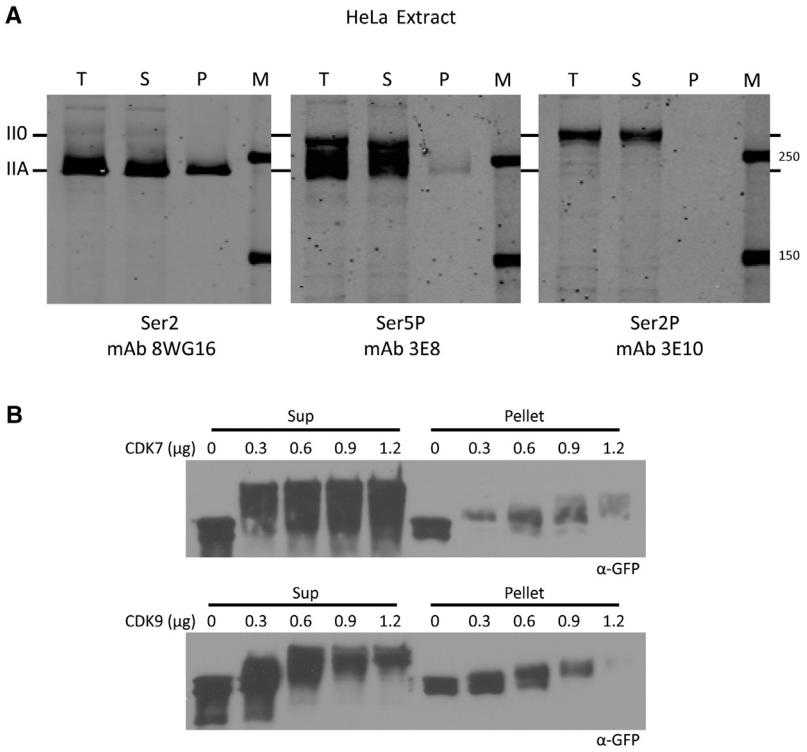

Figure S1. Biotinylated Isoxazole Preferentially Precipitates Unphosphorylated CTD of RNA Polymerase II, Related to Figure 2

(A) Nuclear extracts prepared from cultured HeLa cells were exposed to 100 μM of the biotinylated isoxazole chemical. Total lysate (T), supernatant (S) and pelleted (P) fractions were resolved on SDS-PAGE gels, proteins were visualized by western blotting using antibodies specific to the unphosphorylated form of CTD of RNA polymerase II (left), the serine 5 phosphorylated form of the CTD (middle) and the serine 2 phosphorylated form of the CTD (right). II0: phosphorylated Pol II, IIA: unphosphorylated Pol II.

(B) After phosphorylation by CDK7 or CDK9, GFP:CTD_{C26} was precipitated by 100 μM b-isox. After washing the pellet, precipitated proteins in either supernatant or pellet were visualized by western blotting using GFP antibodies.

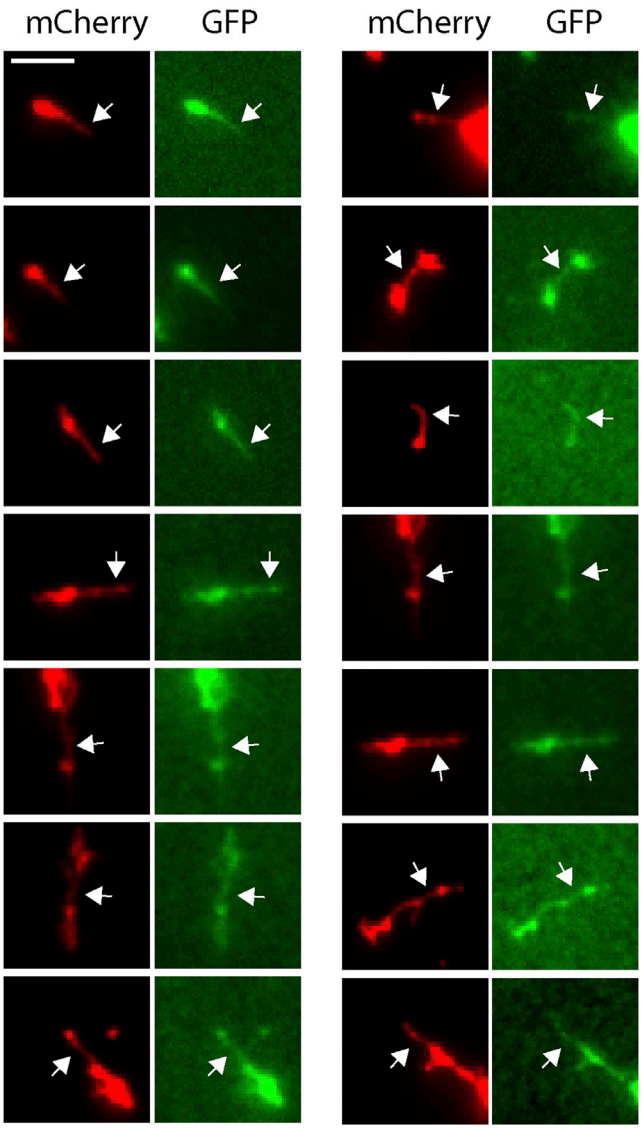

Figure S2. The CTD of RNA Polymerase II Binds to the Lateral Surfaces of TAF15 Polymeric Fibers, Related to Figure 3

Polymeric fibers of mCherry:TAF15 LC domain were mixed with a solution of GFP:CTD_{C26} and immediately visualized by fluorescent microscopy. The TAF15 polymeric fibers form dense tangles that stained strongly with GFP:CTD_{C26}. GFP signals were also detected from single fibers (indicated by arrows), suggesting that GFP:CTD_{C26} binds to the lateral surfaces of TAF15 polymeric fibers. When GFP alone was used as negative control, no GFP signal was detected on the single fibers (data not shown). A scale bar indicates 5 μm.

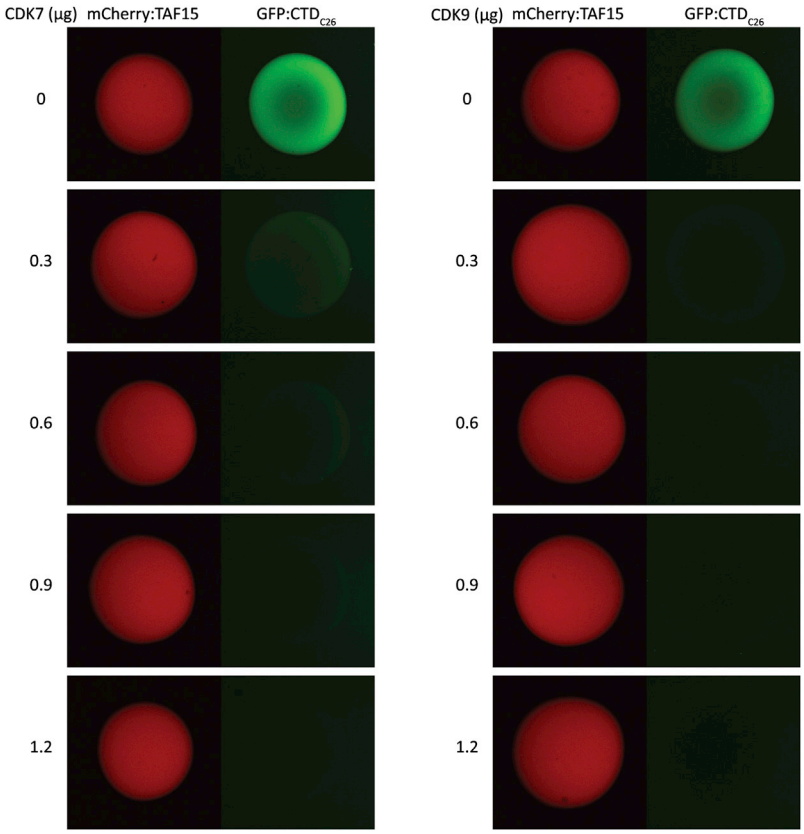

Figure S3. Phosphorylation by CDK7 or CDK9 Inhibits GFP:CTD$_{C26}$ Binding to mCherry:TAF15 Hydrogels, Related to Figure 4

GFP:CTD$_{C26}$ was phosphorylated either by different amounts of either CDK7 (left) or CDK9 (right) for 1 hr at 30°C. Phosphorylated GFP:CTD$_{C26}$ was diluted to 1 μM in a gelation buffer and applied to mCherry:TAF15 hydrogel droplets. After overnight incubation at 4°C, hydrogels were visualized by confocal microscopy (Zeiss LSM510).

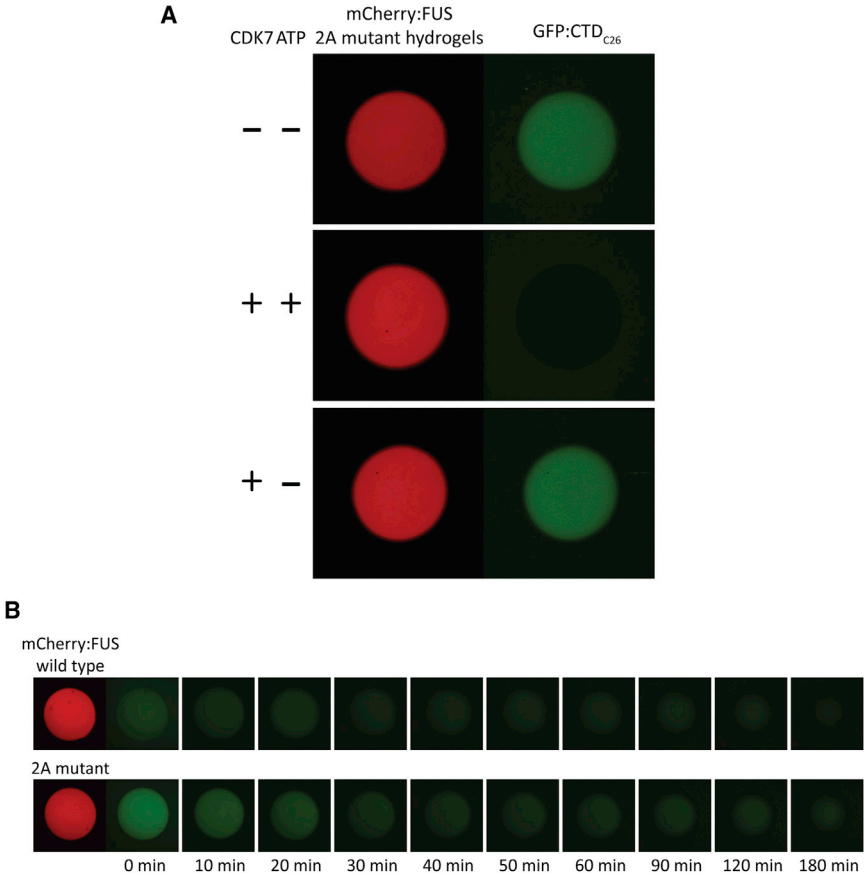

Figure S4. CDK7- and Time-Dependent Release of GFP:CTD$_{C26}$ from Either mCherry:FUS Wild-Type or 2A Mutant Hydrogel Droplets, Related to Figure 5

(A) For phosphorylation, GFP:CTD$_{C26}$ was incubated with or without CDK7 in the presence of ATP for 1 hr at 30°C. GFP:CTD$_{C26}$ was diluted to 1 μM in a gelation buffer and applied to hydrogel droplets formed from mCherry linked to the LC domain of the 2A mutant of FUS (see Figure 1).

(B) GFP:CTD$_{C26}$ was applied to either mCherry:FUS wild-type or 2A mutant hydrogel droplets. After overnight incubation at 4°C, GFP:CTD$_{C26}$ solution was removed and reaction mixtures containing both CDK7 and ATP were applied to the hydrogels. GFP staining to hydrogel droplets was diminished in time-dependent manner as observed by confocal microscopy (Experimental Procedures).

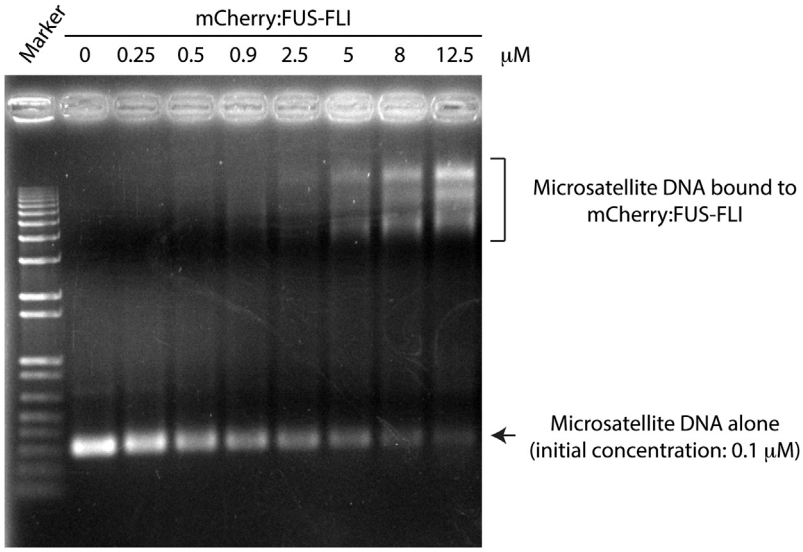

Figure S5. DNA Gel Mobility Shift Assay of mCherry:FUS-FLI Binding to Microsatellite DNA, Related to Figure 6

Microsatellite DNA prepared by PCR amplification from the hNROB1 gene (294 bp, 0.1 μM) was incubated with indicated concentrations of the mCherry:FUS LC domain-FLI DNA-binding domain fusion protein in the presence of 40% glycerol for 1 hr on ice and then separated by electrophoresis on a 1.3% agarose gel.

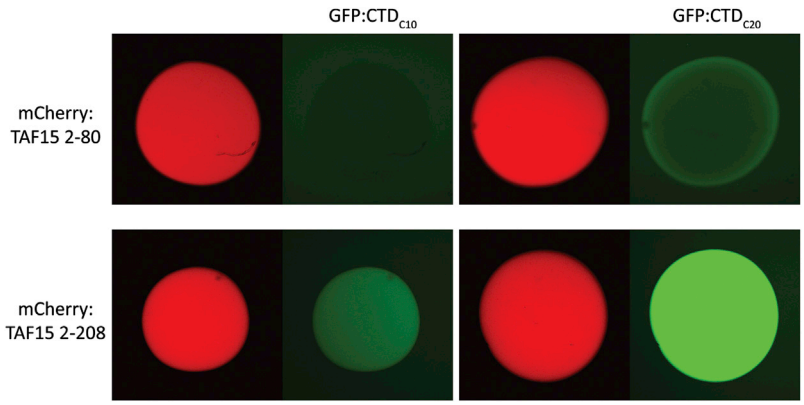

Figure S6. GFP:CTD Binding to Intact or Truncated Version of mCherry:TAF15 Hydrogel Droplets, Related to Figure 7

GFP:CTD$_{C10}$ or GFP:CTD$_{C20}$ were applied to mCherry hydrogel droplets of either intact (residues 2-208) or amino terminal (residues 2-80) of the LC domain of TAF15. After overnight incubation at 4°C, hydrogel droplets were detected by confocal microscopy (Experimental Procedures).