Title

Nuclear protein kinase CLK1 uses a non-traditional docking mechanism to select physiological substrates.

Journal

Biochemical Journal, 472(3)

ISSN

0264-6021

Authors

Keshwani, Malik M
Hailey, Kendra L
Aubol, Brandon E
et al.

Publication Date

2015-12-15

DOI

10.1042/bj20150903

Peer reviewed

HHS Public Access

Author manuscript

Biochem J. Author manuscript; available in PMC 2016 December 15.

Published in final edited form as:

Biochem J. 2015 December 15; 472(3): 329–338. doi:10.1042/BJ20150903.

Nuclear Protein Kinase CLK1 Uses A Nontraditional Docking Mechanism To Select Physiological Substrates

Malik M. Keshwani[1], **Kendra L. Hailey**[2], **Brandon E. Aubol**[1], **Laurent Fattet**[1], **Maria L. McGlone**[1], **Patricia A. Jennings**[2], and **Joseph A. Adams**[1,*]

[1]Department of Pharmacology, University of California, San Diego, La Jolla, CA 92093-0636

[2]Department of Chemistry & Biochemistry, University of California, San Diego, La Jolla, CA 92093-0636

Abstract

Phosphorylation-dependent cell communication requires enzymes that specifically recognize key proteins in a sea of similar, competing substrates. The protein kinases achieve this goal by utilizing docking grooves in the kinase domain or heterologous protein adaptors to reduce “off pathway” targeting. We now provide evidence that the nuclear protein kinase CLK1 (Cdc2-like kinase 1) important for splicing regulation departs from these classic paradigms by using a novel self-association mechanism. The disordered N-terminus of CLK1 induces oligomerization, a necessary event for targeting its physiological substrates, the SR (arginine-serine-rich) protein family of splicing factors. Increasing the CLK1 concentration enhances phosphorylation of the splicing regulator SRSF1 compared to the general substrate myelin basic protein. In contrast, removal of the N-terminus or dilution of CLK1 induces monomer formation and reverses this specificity. CLK1 self-association also occurs in the nucleus, is induced by the N-terminus and is important for localization of the kinase in subnuclear compartments known as speckles. These findings present a new picture of substrate recognition for a protein kinase in which an intrinsically disordered domain is used to capture physiological targets with similar disordered domains in a large oligomeric complex while discriminating against nonphysiological targets.

Keywords

docking; kinase; kinetics; oligomerization; phosphorylation; splicing

[*]To whom correspondence should be sent: Joseph A. Adams, Tel: 858-822-3360, Fax: 858-822-3361, j2adams@ucsd.edu.

Author Contribution

Malik Keshwani expressed and purified hCLK1 and performed substrate specificity, DLS, immunoprecipitation and confocal microscopy experiments. Malik Keshwani also developed the hypothesis of switching substrate specificity through oligomerization. Kendra Hailey performed the SEC experiments. Laurent Fattet helped with confocal imaging. Brandon Aubol performed the concentration-dependent activation experiments for the kinase. Maria McGlone subcloned the hCLK1 gene into pFastBacHTC and generated bacmids. Patricia Jennings provided help with analysis of the CLK1 oligomer and edited the manuscript. Joseph Adams planned the experiments and wrote the paper.

Introduction

Protein phosphorylation regulates countless cellular processes including the fundamental conversion of genetic information into functioning proteins. In the latter, a large precursor mRNA transcript (pre-mRNA) is reduced to a smaller translatable form through splicing, a process whereby introns are removed and exons are selectively joined. Multisite phosphorylation of a family of splicing regulators known as SR proteins is essential for controlling this process [1, 2]. SR proteins contain one or two RNA recognition motifs (RRMs) that bind pre-mRNA and a signature C-terminal RS domain composed of numerous Arg-Ser and several Ser-Pro dipeptides that regulates this interaction along with other steps in RNA processing. Two protein kinase families (SRPKs and CLKs) have been shown to phosphorylate these dipeptides and regulate SR protein function. The SRPKs are located in both the cytoplasm and nucleus where they exclusively phosphorylate Arg-Ser dipeptides in the RS domain [3]. While the cytoplasmic pools are important for SR protein translocation into the nucleus, the function of nuclear SRPKs is not well understood although they have been found attached to components of the spliceosome, a macromolecular RNA-protein machine that incorporates the SR proteins and catalyzes pre-mRNA splicing [4]. In comparison, the CLKs are exclusively localized to the nucleus, and in addition to Arg-Ser dipeptides, uniquely phosphorylate Ser-Pro dipeptides in SR proteins [5, 6]. The latter modification induces conformational changes that shift SR proteins from speckles (interchromatin granules) to sites of transcription and splicing in the nucleus [7]. These specialized functions result in unique CLK-dependent changes in alternative gene splicing [8, 9].

Although SRPKs and CLKs display over-lapping substrate specificities, they use very different recognition mechanisms. SRPK1 possesses a classic docking groove in the kinase domain that initially binds Arg-Ser dipeptides in the RS domain of SRSF1 before transferring them to the active site [10]. Close apposition of the active site and docking groove results in a stable kinase-substrate complex (K_d 100 nM) where the RS domain is efficiently phosphorylated in a directional manner [11]. The CLKs comprise a group of four protein kinases (CLK1-4) that fall into the LAMMER family owing to the conserved sequence 'EHLAMMERILGPLP' in their kinase domains [12]. Although CLKs lack a docking groove, mouse CLK1 (mCLK1) binds SRSF1 with 10-fold greater affinity than SRPK1. Recently, we showed that mCLK1 uses its N-terminus to bind the RS domain and feed dipeptides into the active site [13]. The N-terminus of mCLK1 (approx 140 residues), possessing high sequence redundancy and many disorder-promoting residues, is intrinsically disordered based on structure prediction software [13]. Nonetheless, this extension serves essential catalytic functions by enhancing SR protein affinity (30-fold) and increasing the total number of phosphates added to SRSF1 along with several other SR and SR-like proteins [13]. The mCLK1 N-terminus is flexible and can interact with either the RS domain or the kinase domain. In addition to its catalytic functions, the N-terminus serves as a nuclear localization sequence [14]. Given that CLKs shed the classic docking groove in a well-folded domain, it remains to be seen how a disordered N-terminus can supplant the need for such a groove and achieve efficient phosphorylation of natural substrates in the nucleus.

Outside the conserved kinase domain, CLKs possess structural features that resemble those of their target substrates, a similarity that may provide some clues to their recognition mechanism. Most notably, the N-termini of CLKs are similar in residue content to the RS domains in SR proteins [12]. Like the CLK N-termini, RS domains are thought to be highly disordered, a prediction that was recently confirmed by NMR studies on the unphosphorylated RS domain of SRSF1 [15]. SR proteins can also self-associate and form hetero-complexes with other SR proteins and spliceosomal components through their RS domains [16–19]. These characteristics are likely to be valuable within the congested spliceosomal environment where numerous protein-protein and protein-RNA contacts are made [20, 21]. Furthermore, the ability of RS domains to form protein-protein interactions raises the possibility that CLKs might also self-associate through their N-termini. Duncan et. al. showed that a purified GST-tagged CLK1 can pull-down a Myc-tagged CLK1 as well as a truncated form lacking the kinase domain from cell lysates suggesting that CLK1 may form dimers or possibly larger oligomers in solution through its N-terminus [14]. Although several crystal structures of CLK kinase domains have been solved [22], there is no structure of a full-length CLK that can provide insights into the role of the N-terminus for protein substrate recognition or potential self-assembly mechanisms.

In this new study we investigated the mechanism whereby CLK1 recognizes and phosphorylates SR proteins. We found that the N-terminus of human CLK1 (hCLK1) induces oligomerization and targets the kinase towards its physiological substrate SRSF1. We found that the specific activity of hCLK1 towards SRSF1 increases as a function of total enzyme concentration. At low concentrations, hCLK1 phosphorylates SRSF1 at a similar rate constant as a kinase form lacking the N-terminus, hCLK1_N. However, at higher concentrations, the full-length kinase is, at least, 10-fold faster than hCLK1_N. These findings suggest that hCLK1 undergoes a concentration-dependent change in quaternary structure. Using size exclusion chromatography and dynamic light scattering we showed that hCLK1 is a large oligomer and that removal of the N-terminus results in a monomeric form of the kinase. Since this monomer possesses similar activity as hCLK1_N, oligomerization of the full-length kinase drives substrate specificity. Indeed, we showed that both hCLK1ΔN and the monomeric form of hCLK1 phosphorylate the generic substrate myelin basic protein (MBP) much better than the physiological substrate SRSF1. Thus, the addition of the N-terminus and subsequent oligomerization shifts the substrate specificity of the CLK1 kinase domain from non-physiological to physiological substrates. Furthermore, the N-terminus supports self-association in HeLa cells and directs CLK1 to nuclear speckles. These studies establish a new mechanism in which a protein kinase uses a self-association mechanism to generate a platform for the specific recognition and phosphorylation of physiological targets over generic substrates.

Materials and Methods

Materials

ATP, Mops, Tris, $MgCl_2$, NaCl, EDTA, glycerol, sucrose, acetic acid, lysozyme, DNAse, RNAse, Phenix imaging film, BSA, Protein G–agarose, Ni-resin and liquid scintillant were obtained from Fisher Scientific. ^{32}P-ATP was obtained from NEN Products. Protease

inhibitor cocktail was obtained from Roche. Anti-His monoclonal antibody was purchased from BioLegend. InstantBlue was purchased from Expedeon, Hybond ECL nitrocellulose blotting membrane was purchased from Amersham and the KinaseMax™ Kit was purchased from Ambion.

Expression and Purification of Proteins

Human SRPK1, human SRSF1, mouse CLK1 and mouse CLK1(_N) were expressed in *E. coli* from pET19b vectors containing an N-terminal His_6 tag and purified as previously described [13]. Human kdSRPK1 contains a lysine-to-methionine mutation in the active site and was expressed in *E. coli* from a pET19b vector and purified as previously described [23]. The human CLK1 gene was sub-cloned into pFastBacHTC using Bac-to-Bac Baculoviruses Expression System (Invitrogen) via the BamH1 and Hind III restriction enzyme sites resulting in His_6-CLK1/pFastBacHTC. The new positive constructs were then transformed into *E. coli* DHBAC10 cells to generate the bacmids with the CLK1 gene incorporated. The baculovirus was generated by transfecting the bacmid (Life Technologies protocol). CLK1 virus was transfected and protein expressed in Hi5 insect cells at cell density of $2x10^6$/ml for 48 h. The cells were pelleted and re-suspended in lysis buffer (50mM Tris pH 7.5, 150mM NaCl, 1mM DTT, protease inhibitor cocktail (Roche), 0.75 mM sodium vanadate, 1.25mM sodium pyrophosphate, 0.025% Triton) and stored overnight at -80°C. The cell suspension was thawed at room temperature and sonicated at 25% power for 4.5 min with 15 sec on/off cycles. The lysates were then incubated with the Nickel resin for 1 hr at room temperature followed by 4 sequential block washes (50 ml) of lysis buffer containing 0, 5 and 30 mM imidazole and finally 300 mM NaCL. Elution was performed with 15 ml of 50 mM Tris pH 7.5, 150 mM NaCl, 300 mM imidazole. The eluted protein was dialyzed twice against 50 mM Tris pH 7.5, 150 mM NaCl, 1 mM DTT and 10% Glycerol and was frozen at -80°C.

Phosphorylation Reactions

Phosphorylation reactions were carried out in the presence of MBP or SRSF1, 100 mM Mops (pH 7.4), 10 mM free Mg^{2+}, 5 mg/mL BSA, and 100 μM ^{32}P-ATP (4000–8000 cpm $pmol^{-1}$) at 23 °C for mCLK1 and 37 °C for hCLK1, with the reactions initiated by the addition of enzyme. Single turnover experiments were carried out in the presence of 100 mM Mops (pH 7.4), 10 mM Mg^{2+} and 5 mg/ml BSA, 1 μM enzyme, 200 nM SRSF1 and 100 μM ^{32}P-ATP (4000–8000 cpm/pmol) at 37 °C for hCLK1 and 25 °C for SRPK1, with the reactions initiated by the addition of enzyme. Enzyme concentration-dependent reactions using hCLK1 or hCLK1ΔN (20–750 nM) were carried out by diluting enzyme stocks into 100 mM Mops (pH 7.4), 10 mM free Mg^{2+}, 5 mg/mL BSA, 1 μM SRSF1, and 100 μM ^{32}P-ATP (4000–8000 cpm $pmol^{-1}$) at 37 °C and allowing the reaction to proceed for 2 or 4 minutes. All reactions were carried out in a total volume of 10 μL and quenched with 10 μL SDS-PAGE loading buffer. Phosphorylated SRSF1 and MBP were separated from unreacted ^{32}P-ATP on a 10% SDS-PAGE by running for 1 hour at 170V. Protein bands were cut from the dried gel and quantitated on the ^{32}P channel in liquid scintillant. The total amount of phosphoproduct was then determined by considering the specific activity (cpm/min) of the reaction mixture.

Size Exclusion Chromatography & Dynamic Light Scattering

Affinity-purified and concentrated hCLK1 (15 μM) and hCLK1ΔN (30 μM) were run on a S300 gel filtration column using a Biorad FPLC system in a buffer containing 50mM Tris (pH 7.5), 150 mM NaCl, 1 mM DTT and 5% Glycerol. Molecular weight standards were also run under the same conditions. The DLS instrument (Protein Solutions DynaPro) was calibrated with the gel filtration buffer and light scattering experiments were performed on purified hCLK1 (1.5μM and 5μM) and hCLK1ΔN (20 μM). Bovine serum albumin was used as a control.

Western Blot Analyses

HeLa cells were plated overnight to achieve 60% confluence and transfected with RFP-tagged CLK1 constructs (5_g) for 24 h along with a mock control. The cells were harvested and lysed using 10 mM Hepes (pH 7.9), 1.5 mM $MgCl_2$, 10 mM KCl, 1 mM DTT, 0.05% Triton and protease inhibitor cocktail. The lysates were centrifuged at 250 g for 10 min to pellet the nucleus and the supernatant was retained as the cytoplasmic fraction. The pellet was washed with lysis buffer and re-suspended in 200 _L of a solution containing 0.25 mM sucrose and 10 mM $MgCl_2$. The nuclear suspension was layered on a bed of 0.88 mM sucrose and 0.5 mM $MgCl_2$ and spun at 1100g to obtain a clean nuclear pellet. The pellet was re-suspended in 300_L of 1X RIPA buffer (Cell signaling), spun at 1100g to pellet any solids and the supernatant was retained as nuclear fractions. The nuclear fractions were concentration-matched using Bradford assays at 1mg/mL and 100_L of lysate was used for the immunoprecipitation reaction. The lysates were pre-cleared with 15_L of protein G beads and spun at 1609g for 3 min. Rabbit poly-clonal RFP antibody (2_L, abcam) was added to the mock-treated and CLK1 over-expressed nuclear lysates along with 25_L slurry of Protein G beads (previously washed 3X in wash buffer 50mM Tris pH 7.3, 150mM NaCl and 0.1% NP40). The lysates were incubated on a nutator overnight in a cold room. The immunoprecipitated beads were washed 3X with 500_L of wash buffer and collected by spinning at 1609g for 3 min. 2x SDS buffer (40_L) was added and the beads were heated at 100 ºC for 5 min followed by spinning at 17000g to precipitate the beads. The supernatant (20_L each) was resolved on 9% SDS-PAGE gel follow by western blot analysis and assessed for the presence of endogenous CLK1. Nuclear lysates (10_L) were used for assessing overexpression of RFP-tagged CLK1 using RFP antibody as well as SRSF1 as a marker for nuclear fractions.

Live Cell Imaging Using Confocal Microscopy

HeLa cells were seeded and transfected in a 35-mm poly-D-lysine coated plate (Mattek). When 60% confluent, cells were transfected separately for 24h with various RFP-tagged forms of hCLK1 (1.5μg) using Fugene transfection reagent. Live cells were then imaged for RFP using an Olympus FV1000 with a 488-laser line. Cells were also co-transfected with RFP-tagged forms of hCLK1 and GFP-SRSF1 (1μg each) for 24h using Fugene transfection reagent and then imaged for both RFP and GFP as described above. All images were linearly analysed and pseudo-colored using ImageJ analysis software.

Results & Discussion

N-Terminus of mCLK1 Regulates Substrate Specificity

Previous studies from our lab indicate that SR protein recognition and high-level, multisite phosphorylation is dependent on the N-terminus in mCLK1 (aka Clk/Sty) [13]. We wished to ask whether this activation phenomenon is specific to SR proteins or could be replicated using other substrates. For this, we choose the general LAMMER kinase substrate MBP (myelin basic protein). Although MBP lacks an RS domain, previous mapping experiments showed that mCLK1 readily phosphorylates one Ser-Pro dipeptide (GSPMAR) in this protein substrate [24]. We investigated MBP phosphorylation using mCLK1, previously expressed in our lab as a His-tagged kinase in *E. coli* [6], and compared it to a form lacking the N-terminus, a 134-residue extension (mCLK1ΔN) (Fig. 1A). The phosphorylation of SRSF1 and MBP was monitored using equal concentrations of mCLK1 or mCLK1ΔN (Fig. 1B). Time-dependent phosphate incorporation was plotted to determine reaction velocities (Fig. 1 C,D). While the initial velocities for MBP and SRSF1 are similar using mCLK1ΔN, the initial velocity for SRSF1 phosphorylation is about 10-fold higher than that for MBP phosphorylation using wild-type mCLK1 (Fig. 1C,D). We also explored an additional substrate lacking an RS domain and found that while histone H1 is a poor substrate for mCLK1 compared to SRSF1 and MBP, the N-terminus likewise did not significantly enhance its phosphorylation rate (Fig. 1E). Overall, the data show that the N-terminus of mCLK1 plays a vital kinetic role in targeting the kinase domain toward SR proteins as opposed to substrates lacking RS domains.

Unique CLK1 Substrate Specificity Is Not Species Dependent

Owing to the special role of the N-terminus for SR protein phosphorylation, we wished to learn more about the structure of the full-length CLK1. Since we were unable to purify large amounts of mCLK1, we explored the possibility of expressing human CLK1 (hCLK1). Unfortunately, we were unable to express this kinase in *E. coli* despite similarities between the mouse and human enzymes (80% sequence identity). However, using a baculovirus system, we were able to obtain approximately 3mg of pure hCLK1 per liter of growth medium (Fig. 2A). The purified hCLK1 induced the classic gel shift for SRSF1 on SDS-PAGE consistent with Ser-Pro phosphorylation (Fig. 2B). Using single turnover experiments, we found that hCLK1 adds about 18–19 phosphates onto SRSF1, consistent with prior findings for the mouse enzyme [5, 6] (Fig. 2C). As a control, we showed that SRPK1 adds 15 phosphates onto SRSF1 which includes all the available Arg-Ser dipeptides in the RS domain confirming the viability of the substrate in our assays (Fig. 2C). To determine whether hCLK1 displays similar substrate specificity as mCLK1, we expressed in *E. coli* and purified a form lacking the N-terminus, hCLK1ΔN (Fig. 2A). We found that hCLK1 showed higher catalytic activity toward SRSF1 than MBP in keeping with the mouse enzyme (Fig. 2D). In contrast, hCLK1ΔN phosphorylated MBP more efficiently than SRSF1 (Fig. 2D). Thus, the N-terminus in hCLK1 appears to play a more significant role in regulating specificity compared to mCLK1. The intrinsic differences in SRSF1/MBP specificities of the human and mouse kinases in the absence of their N-termini are difficult to decipher at this time. Although the kinase domains share 95% sequence identity, there are several unique charged residues that may account for these specificity variations. Overall,

these studies demonstrate that the substrate specificity of the kinase is not species-dependent and is generally induced by the N-terminus of CLK1.

N-Terminus Induces Oligomerization of hCLK1

Having expressed pure hCLK1 at high levels, we next wished to obtain structural information on this protein. Although several crystal structures for the kinase domain have been published there is currently no structure for any full-length form [22]. Given a prior report suggesting that CLK1 may form dimers [12], we initially subjected both hCLK1 and hCLK1ΔN to size exclusion chromatography (SEC) and found that they eluted in different fractions. hCLK1ΔN eluted as a single peak with a retention time consistent with a 39 kDa monomer based on a panel of molecular weight standards (Fig. 3A). In comparison, hCLK1 eluted much earlier in the SEC chromatogram, just following blue dextran, a standard that defines the column void volume (Fig. 3A). Thus, whereas the kinase domain runs as a monomer, the full-length kinase runs as a large oligomer with a molecular mass of, at least, 1000 kDa. To provide additional support for these findings we performed dynamic light scattering (DLS) on both proteins. We found that hCLK1ΔN (20 μM) displayed a hydrodynamic radius (4 nm) consistent with a monomer (Fig. 3B). Interestingly, at 13-fold lower concentration, hCLK1 (1.5 μM) displayed a very large hydrodynamic radius of 100 nm by DLS, consistent with a large, self-associated oligomer (Fig. 3B). Unfortunately, we were unable to detect a monomeric form of hCLK1 in the DLS experiments owing to poor signal at low protein concentrations. Increasing the hCLK1 concentration by about 3-fold (5 μM) leads to further increases in particle size (Fig. 3B). Overall, the data indicate that the N-terminus induces a large multimeric conformation of hCLK1 in solution.

Oligomer Formation Activates hCLK1

To evaluate whether oligomer formation affects catalysis, we measured SRSF1 phosphorylation (1 μM) upon successive dilutions of hCLK1 (Fig. 4A). Reaction velocities were normalized to the total amount of kinase in each reaction and then plotted as a function of total hCLK1 concentration (20–750 nM). For a monomeric kinase, the specific activity should remain constant over a broad range of enzyme concentrations. Surprisingly, we found that the specific activity of hCLK1 toward SRSF1 increased by about 7-fold (0.05–0.35 min^{-1}) over the experimental concentration range. Furthermore, it is likely that this is a lower limit on the degree of activation as the hCLK1 specific activity increases throughout the experiment. Since the total substrate is held constant, we performed a control experiment at the highest hCLK1 concentration to make sure that any changes in specific activity are not related to changes in free SRSF1. We showed that at 750 nM hCLK1, increasing the total SRSF1 from 1 to 1.8 μM had no effect on the observed rates (Fig. 4C). These findings indicate that the reduced specific activity at low hCLK1 is not the result of excess SRSF1 binding to an inhibitory site on the kinase. We showed that, unlike the full-length kinase, the specific activity of hCLK1ΔN did not change over a similar concentration range, in keeping with its monomeric form (Fig. 4A). These findings indicate that hCLK1 undergoes a concentration-dependent change in quaternary structure. Given the solution structural data (Fig. 3), we propose that this change reflects a transition from a monomeric to a large oligomeric form. The near convergence of specific activities at low kinase concentrations suggests that hCLK1 is most likely a monomer at these lower levels. Interestingly, the

kinetic data shows no signs of a plateau at high concentrations, consistent with the DLS results where further increases in total hCLK1 led to increased particle sizes (Fig. 3B). Overall, the data show that the ability of the N-terminus to activate CLK1 toward SR proteins is driven largely through oligomerization.

Substrate Specificity Regulation Through the N-Terminus & Oligomerization

Since oligomerization activates hCLK1 toward SRSF1 phosphorylation, we wished to determine whether this phenomenon regulates substrate specificity. We measured the phosphorylation rates of a fixed amount of MBP (13 μM) and found that increasing hCLK1 did not impact the enzyme-normalized initial velocities over a wide range (Fig. 4B). Under conditions where large changes in specific activity are observed for SRSF1 (100–750 nM), no significant changes are observed for MBP phosphorylation. As a control, we measured the phosphorylation of MBP using hCLK1ΔN and found no changes in enzyme-normalized velocities, consistent with a monomeric form of the kinase (Fig. 4B). These findings demonstrate that whereas oligomerization greatly enhances phosphorylation rates for the physiological substrate SRSF1, they have no effect on those for MBP which lacks an RS domain. These studies now allow us to parse the effects of both the N-terminus and oligomerization for substrate specificity. By comparing the specific activities of hCLK1ΔN at any concentration, hCLK1 at the lowest concentration and hCLK1 at the highest concentration, we can define the catalytic parameters of the separate kinase domain, the monomeric, full-length hCLK1 and the oligomeric hCLK1 (Fig. 4D). These comparisons demonstrate that the addition of the N-terminus onto the kinase domain has a significant, negative effect on MBP (↓ 10-fold) and only a minor effect (↑ <2-fold) on SRSF1 phosphorylation. Conversely, the induction of the oligomeric form of hCLK1 at higher concentrations has a significant, positive effect on SRSF1 (↑ >7-fold) and no effect on MBP phosphorylation. By comparing the relative specific activities of these kinase forms with regard to SRSF1 and MBP, we can show that the addition of the N-terminus followed by oligomerization has a compound effect on specificity (Fig. 4E). Whereas the addition of the N-terminus abolishes the strong preference for MBP phosphorylation resulting in a largely non-specific enzyme, oligomerization further switches preferences such that the kinase now strongly prefers its physiological substrate SRSF1 over MBP. Finally, since the extent of oligomerization increases at high hCLK1 levels, the observed specificity change due to oligomerization is only a lower limit and is likely to be much higher. Overall, these studies show that oligomerization through the N-terminus not only activates the kinase toward its natural substrate but also helps to discriminate against unnatural substrates.

Self-Association of hCLK1 in HeLa Cells

Having shown the relevance of the N-terminus for *in vitro* oligomerization and catalytic function, we wished to determine whether hCLK1 self-associates *in vivo*. To address this possibility, we designed an RFP-tagged form of full-length hCLK1 (CLK1-RFP) and expressed it in HeLa cells (Fig. 5A). For these studies we attached RFP to the C-terminal end of the kinase to avoid potential interference with N-terminus-induced oligomerization. We showed that immunoprecipitated CLK1-RFP interacts with endogenous hCLK1 in nuclear lysates (Fig. 5B). Although prior studies showed that an over-expressed, myc-tagged mCLK1 can be pulled down by a recombinant GST-mCLK1 added to cell lysates, our

studies now show that these kinase-kinase interactions are stable at endogenous levels of hCLK1. To determine whether the N-terminus is important for self-association, we expressed a form of hCLK1 lacking the N-terminus but containing a nuclear localization sequences (NLS) from nucleoplasmin 2 (KRLVPQKQASVAKKKK) (Fig. 5A). We showed that this construct, NLS-CLK1ΔN-RFP, localizes in the nucleus similar to CLK1-RFP (Fig. 5C). However, unlike the full-length kinase, NLS-CLK1ΔN-RFP did not interact with endogenous hCLK1 (Fig. 5B). While these studies do not address the extent of oligomerization, they indicate hCLK1 is very capable of self-association in cells through its N-terminus, an observation in line with our *in vitro* data.

Self-Association Regulates hCLK1 Entry into Nuclear Speckles

We showed previously that nuclear speckles disassemble upon CLK1 phosphorylation of Ser-Pro dipeptides in SRSF1 [9]. Since CLK1 has also been shown to reside in speckles with SR proteins, it is possible that this disassembly mechanism may require direct entry of the kinase into speckles. We wished to test this hypothesis and also determine whether the N-terminus could play a role in this mechanism. Although active CLK1 has been reported in speckles, visualization of the inactive kinase is more robust since it does not disturb speckle integrity [12]. Indeed, we showed that a kinase inactive form of hCLK1 (kdCLK1-RFP) localizes clearly in nuclear speckles (Fig. 5C). To evaluate whether the N-terminus plays a role in subnuclear localization, we generated an inactive RFP-tagged form of CLK1 lacking its N-terminus (Fig. 5A). We found that NLS-CLK1ΔN-RFP was diffusely localized in the nucleus similar to CLK1-RFP (Fig. 5C). In comparison, the inactive NLS-kdCLK1ΔN-RFP was also diffuse and showed no prominent speckle patterns as observed for kdCLK1-RFP (Fig. 5C). These findings suggest that the N-terminus is important for entry of hCLK1 into speckles. To determine whether the N-terminus maintains the integrity of the speckles, we co-expressed GFP-tagged SRSF1 (GFP-SRSF1) and showed that both CLK1-RFP and NLS-CLK1ΔN-RFP could disperse the speckles (Fig. 5D). Cells that did not express an active kinase, displayed classic speckle formation when monitoring GFP-SRSF1 (Fig. 5D). Overall, these findings demonstrate that while the N-terminus is important for speckle localization, the kinase domain by itself is capable of diffusing speckles.

Nuclear speckles are dynamic, membrane-free structures that act as storage depots for many splicing factors including SR proteins [25]. CLK-dependent phosphorylation mobilizes SR proteins from speckles to sites of transcription/splicing on the periphery of the speckle [7]. Although our data show that the N-terminus of CLK1 is necessary for speckle localization of CLK1, it is not necessary for dissolution of the speckles. Previous fluorescence studies indicate that SR proteins rapidly diffuse in and out of speckles [26]. It is possible that the N-terminus is not required for speckle disassembly in our experiments since the pool of nucleoplasmic kinase can readily phosphorylate SR proteins transiently released from speckles. The advantage of the N-terminus may be kinetic in that the proximity of CLK1 in the speckles could facilitate rapid SR protein phosphorylation, a phenomenon that cannot be observed in our expression experiments. In keeping with this idea, we have observed large rate enhancements for the full-length hCLK1 compared to the hCLK1ΔN under conditions where the kinase is highly oligomerized (Fig. 4A). Given these observations, the function of the N-terminus in the nucleus may be 2-fold. First, the N-terminus may induce a high

activity form of the kinase though self-association. In the speckle, the effective concentration of the kinase may be exceedingly high, thereby facilitating oligomerization and increased catalytic activity. Second, the N-terminus may position the kinase near the SR proteins in storage speckles for rapid proximal phosphorylation.

A New Model for Kinase-Substrate Recognition

Protein kinases use several strategies for attaining efficient substrate phosphorylation. A common theme is the utilization of a binding cavity in a folded domain/subunit that recognizes complementary charged/hydrophobic residues in the protein substrate. For example, SRPKs use an electronegative docking groove in the large lobe of the kinase domain to recognize positively charged Arg-Ser dipeptides in the RS domain of SRSF1. In the cyclin-dependent kinase Cdk2 or the nonreceptor tyrosine kinase Src, noncatalytic domains (i.e.- cyclinA, SH2 & SH3) recognize sequences remote from the phosphorylation site in the target substrate [27, 28]. In contrast to these classic mechanisms, X-ray structures of CLK1 and CLK3 lacking N-termini reveal that no such docking grooves are present in these kinase domains. In a prior study we showed that the N-terminus of mCLK1 interacts with the RS domain of SRSF1 and greatly enhances phosphorylation rates but it is unclear how the N-terminus facilitates catalysis [13]. In this new study we demonstrate that this interaction leads to catalytic activation within the context of an hCLK1 oligomer. By comparing the phosphorylation rates of its physiological target with that for the generic substrate MBP, we can now define how the N-terminus selects substrates (Fig. 6). Although the kinase domain has a strong preference for phosphorylating MBP over SRSF1, the introduction of the N-terminus abolishes this preference by raising the energy barrier for MBP without greatly affecting that for SRSF1 (Fig. 6(a)). This specificity change could result from the N-terminus blocking non-RS domain containing substrates. At higher concentrations, the attachment of the N-terminus induces oligomerization of hCLK1, a change in quaternary structure that lowers the energy barrier for SRSF1 without affecting that for MBP (Fig. 6(b)). The net effect of the addition of the N-terminus and then subsequent oligomerization is a demonstrative switch in substrate specificity from nonphysiological to physiological substrates. How the oligomer enforces this specificity switch is not fully understood at this time but we propose that overlapping interactions between multiple, unfolded N-termini in the complex favor highly productive contacts with the similar, unfolded RS domains of the SR proteins. This new phosphorylation mechanism for CLK1 departs from the traditional kinase-substrate recognition mechanisms in which folded structural domains are employed to recognize segments of the protein substrates. In this new model, kinase-substrate recognition is facilitated by complex networking between two unstructured domains in the enzyme and its protein target.

Acknowledgments

Funding Sources

This work was supported by an NIH grant (GM98528) to J.A.A.

Abbreviations

CLK1	Cdc2-like kinase 1
RRM	RNA recognition motif
RS domain	domain rich in arginine-serine repeats
SR protein	splicing factor containing a C-terminal RS domain
SRPK1	SR-specific protein kinase 1
SRSF1	SR protein splicing factor 1 (aka ASF/SF2)

References

1. Reed R. Initial splice-site recognition and pairing during pre-mRNA splicing. Curr Opin Genet Dev. 1996; 6:215–220. [PubMed: 8722179]
2. Mermoud JE, Cohen P, Lamond AI. Ser/Thr-specific protein phosphatases are required for both catalytic steps of pre-mRNA splicing. Nucleic acids research. 1992; 20:5263–5269. [PubMed: 1331983]
3. Wang HY, Lin W, Dyck JA, Yeakley JM, Songyang Z, Cantley LC, Fu XD. SRPK2: a differentially expressed SR protein-specific kinase involved in mediating the interaction and localization of pre-mRNA splicing factors in mammalian cells. The Journal of cell biology. 1998; 140:737–750. [PubMed: 9472028]
4. Mathew R, Hartmuth K, Mohlmann S, Urlaub H, Ficner R, Luhrmann R. Phosphorylation of human PRP28 by SRPK2 is required for integration of the U4/U6-U5 tri-snRNP into the spliceosome. Nature structural & molecular biology. 2008; 15:435–443.
5. Aubol BE, Plocinik RM, Hagopian JC, Ma CT, McGlone ML, Bandyopadhyay R, Fu XD, Adams JA. Partitioning RS Domain Phosphorylation in an SR Protein through the CLK and SRPK Protein Kinases. Journal of molecular biology. 2013
6. Velazquez-Dones A, Hagopian JC, Ma CT, Zhong XY, Zhou H, Ghosh G, Fu XD, Adams JA. Mass spectrometric and kinetic analysis of ASF/SF2 phosphorylation by SRPK1 and Clk/Sty. The Journal of biological chemistry. 2005; 280:41761–41768. [PubMed: 16223727]
7. Sacco-Bubulya P, Spector DL. Disassembly of interchromatin granule clusters alters the coordination of transcription and pre-mRNA splicing. The Journal of cell biology. 2002; 156:425–436. [PubMed: 11827980]
8. Yomoda J, Muraki M, Kataoka N, Hosoya T, Suzuki M, Hagiwara M, Kimura H. Combination of Clk family kinase and SRp75 modulates alternative splicing of Adenovirus E1A. Genes Cells. 2008; 13:233–244. [PubMed: 18298798]
9. Keshwani MM, Aubol BE, Fattet L, Ma CT, Qiu J, Jennings PA, Fu XD, Adams JA. Conserved proline-directed phosphorylation regulates SR protein conformation and splicing function. The Biochemical journal. 2015; 466:311–322. [PubMed: 25529026]
10. Ngo JC, Giang K, Chakrabarti S, Ma CT, Huynh N, Hagopian JC, Dorrestein PC, Fu XD, Adams JA, Ghosh G. A sliding docking interaction is essential for sequential and processive phosphorylation of an SR protein by SRPK1. Molecular cell. 2008; 29:563–576. [PubMed: 18342604]
11. Ma CT, Velazquez-Dones A, Hagopian JC, Ghosh G, Fu XD, Adams JA. Ordered multi-site phosphorylation of the splicing factor ASF/SF2 by SRPK1. Journal of molecular biology. 2008; 376:55–68. [PubMed: 18155240]
12. Colwill K, Pawson T, Andrews B, Prasad J, Manley JL, Bell JC, Duncan PI. The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. The EMBO journal. 1996; 15:265–275. [PubMed: 8617202]

13. Aubol BE, Plocinik RM, Keshwani MM, McGlone ML, Hagopian JC, Ghosh G, Fu XD, Adams JA. N-terminus of the protein kinase CLK1 induces SR protein hyperphosphorylation. The Biochemical journal. 2014; 462:143–152. [PubMed: 24869919]
14. Duncan PI, Howell BW, Marius RM, Drmanic S, Douville EM, Bell JC. Alternative splicing of STY, a nuclear dual specificity kinase. The Journal of biological chemistry. 1995; 270:21524–21531. [PubMed: 7665564]
15. Xiang S, Gapsys V, Kim HY, Bessonov S, Hsiao HH, Mohlmann S, Klaukien V, Ficner R, Becker S, Urlaub H, Luhrmann R, de Groot B, Zweckstetter M. Phosphorylation drives a dynamic switch in serine/arginine-rich proteins. Structure. 2013; 21:2162–2174. [PubMed: 24183573]
16. Xiao SH, Manley JL. Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. Genes & development. 1997; 11:334–344. [PubMed: 9030686]
17. Mermoud JE, Cohen PT, Lamond AI. Regulation of mammalian spliceosome assembly by a protein phosphorylation mechanism. The EMBO journal. 1994; 13:5679–5688. [PubMed: 7988565]
18. Xiao SH, Manley JL. Phosphorylation-dephosphorylation differentially affects activities of splicing factor ASF/SF2. The EMBO journal. 1998; 17:6359–6367. [PubMed: 9799243]
19. Jiang Z, Tang H, Havlioglu N, Zhang X, Stamm S, Yan R, Wu JY. Mutations in tau gene exon 10 associated with FTDP-17 alter the activity of an exonic splicing enhancer to interact with Tra2 beta. The Journal of biological chemistry. 2003; 278:18997–19007. [PubMed: 12649279]
20. Jurica MS, Moore MJ. Pre-mRNA splicing: awash in a sea of proteins. Molecular cell. 2003; 12:5–14. [PubMed: 12887888]
21. Jurica MS. Detailed close-ups and the big picture of spliceosomes. Current opinion in structural biology. 2008; 18:315–320. [PubMed: 18550358]
22. Bullock AN, Das S, Debreczeni JE, Rellos P, Fedorov O, Niesen FH, Guo K, Papagrigoriou E, Amos AL, Cho S, Turk BE, Ghosh G, Knapp S. Kinase domain insertions define distinct roles of CLK kinases in SR protein phosphorylation. Structure. 2009; 17:352–362. [PubMed: 19278650]
23. Yeakley JM, Tronchere H, Olesen J, Dyck JA, Wang HY, Fu XD. Phosphorylation regulates in vivo interaction and molecular targeting of serine/arginine-rich pre-mRNA splicing factors. The Journal of cell biology. 1999; 145:447–455. [PubMed: 10225947]
24. Menegay HJ, Myers MP, Moeslein FM, Landreth GE. Biochemical characterization and localization of the dual specificity kinase CLK1. Journal of cell science. 2000; 113(Pt 18):3241–3253. [PubMed: 10954422]
25. Spector DL, Lamond AI. Nuclear speckles. Cold Spring Harbor perspectives in biology. 2011; 3
26. Misteli T, Caceres JF, Spector DL. The dynamics of a pre-mRNA splicing factor in living cells. Nature. 1997; 387:523–527. [PubMed: 9168118]
27. Goldsmith EJ, Akella R, Min X, Zhou T, Humphreys JM. Substrate and docking interactions in serine/threonine protein kinases. Chem Rev. 2007; 107:5065–5081. [PubMed: 17949044]
28. Brown NR, Noble ME, Endicott JA, Johnson LN. The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. Nature cell biology. 1999; 1:438–443. [PubMed: 10559988]

Author Manuscript

Author Manuscript

Author Manuscript

Summary Statement

CLK1 uses an oligomerization mechanism to recognize its physiological protein substrates.

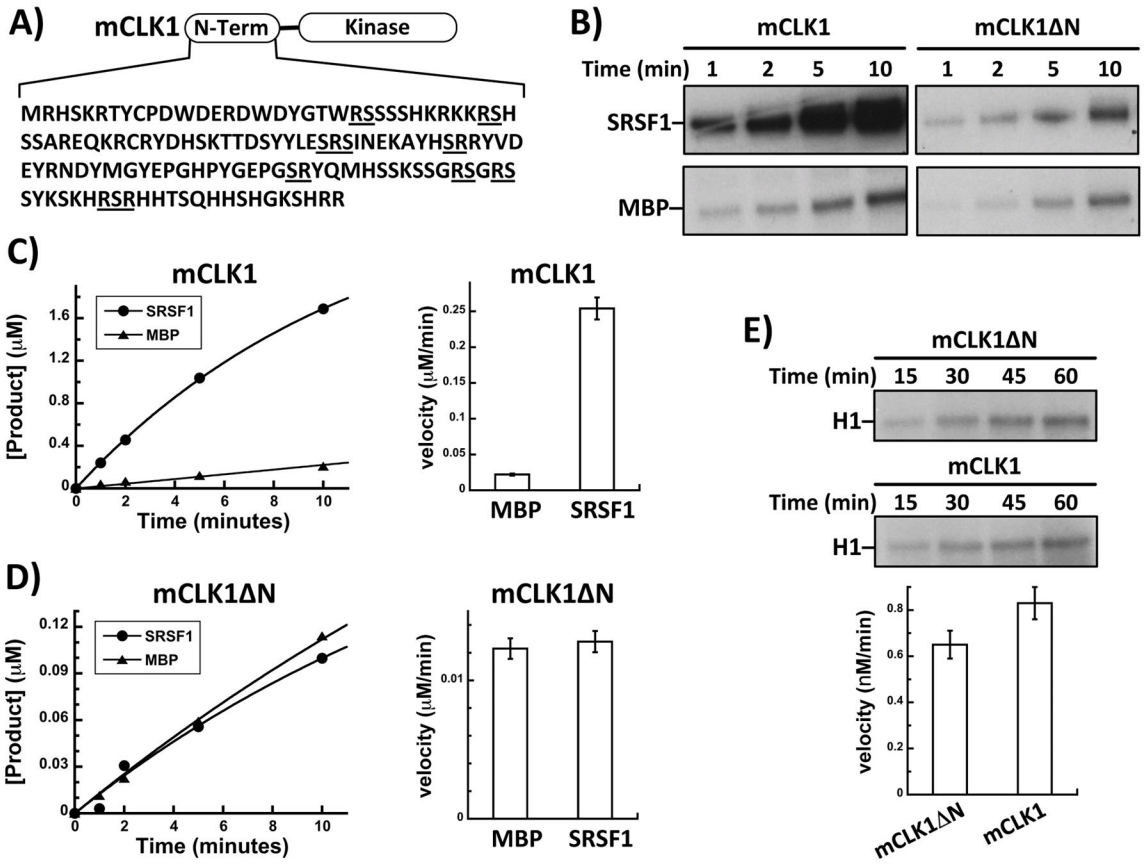

Figure 1. N-terminus of mCLK1 Impacts SR Protein Specificity

A) mCLK1 structure. RS/SR dipeptides are underlined. B) Autoradiograms of SRSF1 and MBP phosphorylation by mCLK1 and mCLK1ΔN. Reactions contain 500 nM mCLK1 or mCLK1ΔN and 1 μM MBP or SRSF1. C,D) Progress curves (left panels) and initial velocity measurements (right panels) for SRSF1 and MBP phosphorylation using mCLK1 (C) and mCLK1ΔN (D). E) Autoradiogram and initial velocity measurements for histone H1 phosphorylation by mCLK and mCLK1ΔN.

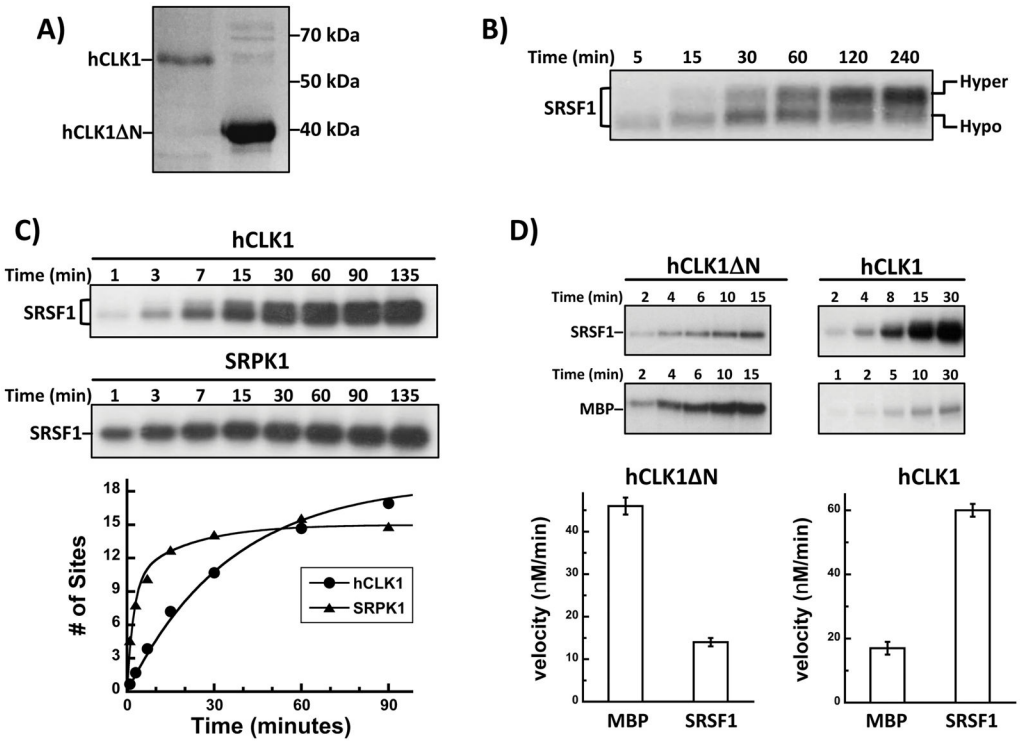

Figure 2. Kinetic Characterization of hCLK1

A) Purified forms of hCLK1 expressed in Sf9 insect cells and hCLK1ΔN expressed in *E. coli*. B) Autoradiogram of SRSF1 phosphorylation by hCLK1 showing a gel shift. C) Single turnover progress curves for SRSF1 phosphorylation using hCLK1 and SRPK1. Concentrations of kinases and SRSF1 are 1 μM and 0.2 μM. The data for hCLK1 were fit to a rate constant and maximum amplitude of 0.3 min^{-1} and 19 sites. The data for SRPK1 were fit to a double exponential with rate constants and amplitudes of 0.4 and 0.05 min^{-1} and 10 and 5 sites. D) Phosphorylation progress curves for SRSF1 and MBP using hCLK1 and hCLK1ΔN. Reactions contain 300 nM hCLK1 or hCLK1ΔN and 1 μM MBP or SRSF1.

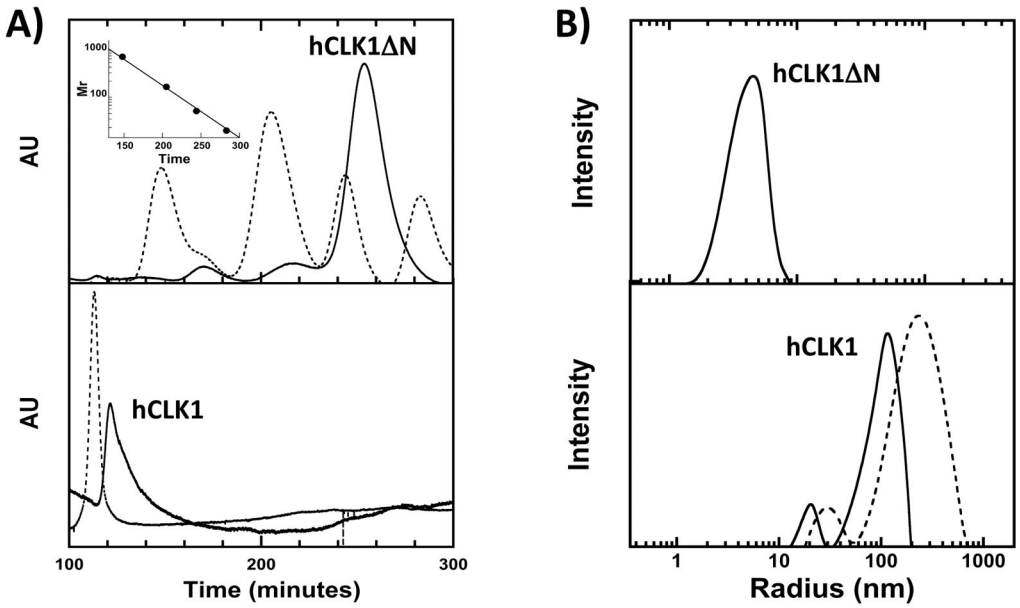

Figure 3. Structural Analyses of hCLK1 and hCLK1ΔN

A) Size exclusion chromatography for hCLK1 (lower panel) and hCLK1ΔN (upper panel). Both proteins are run on a Sephacryl S-300 column (solid line) and eluted with a buffer containing 50 mM Mops (pH 7), 150 mM NaCl, 2 mM DTT and 5% glycerol. Standard proteins were run (dotted line) and their molecular masses were plotted against retention time to determine a molecular mass of 39 kDa for hCLK1ΔN (inset). B) Dynamic light scattering for hCLK1 (lower panel) and hCLK1ΔN (upper panel). Data were recorded at 20 μM hCLK1ΔN and 1.5 μM (solid line) and 5 μM (dotted line) for hCLK1.

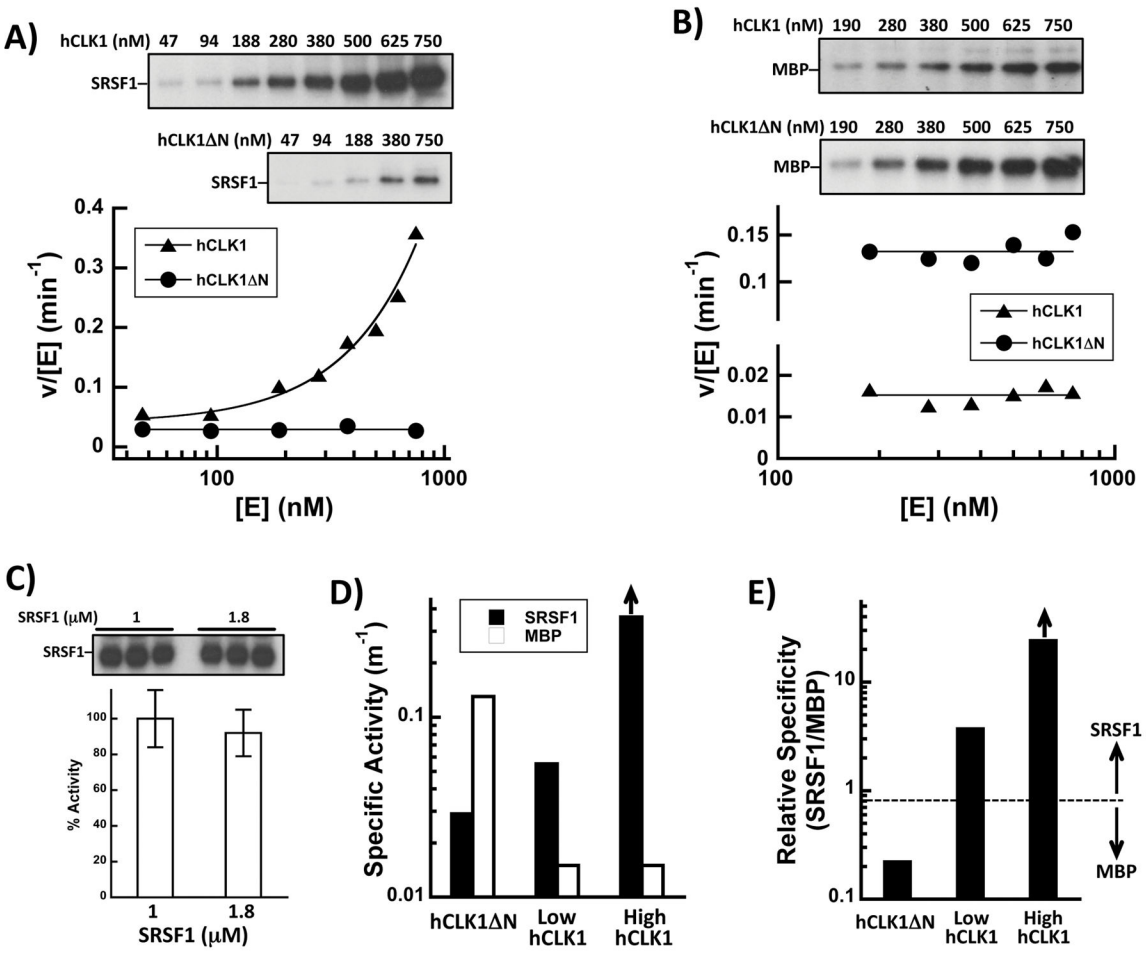

Figure 4. Substrate Specificity As A Function of Oligomerization

A) Concentration-dependent changes in hCLK1 and hCLK1ΔN specific activity using SRSF1. Phosphorylation of 1 μM SRSF1 is assessed after 2-minute incubation with varying hCLK1 or a 4 minute-incubation with varying hCLK1ΔN. The velocities normalized against the total enzyme concentration (specific activity) are plotted as a function of total enzyme concentration. B) Concentration-dependent changes in hCLK1 and hCLK1ΔN specific activities using MBP. Phosphorylation of 13 μM MBP is assessed after 7-minute incubation with varying hCLK1 or hCLK1ΔN. The velocities normalized against the total enzyme concentration (specific activity) are plotted as a function of total enzyme concentration. C) Increasing SRSF1 concentration on reaction velocity. Phosphorylation velocity is monitored in triplicate using 0.75 μM hCLK1 and 1 or 1.8 μM total SRSF1. D) Bar graph showing the specific activities of hCLK1ΔN and hCLK1 (at the lowest and highest concentrations) using SRSF1 and MBP. E) Bar graph showing the relative specific activities of hCLK1ΔN and hCLK1 (at the lowest and highest concentrations) using SRSF1 and MBP. Values >1 imply larger specificity for SRSF1 versus MBP whereas values <1 imply larger specificity for MBP versus SRSF1.

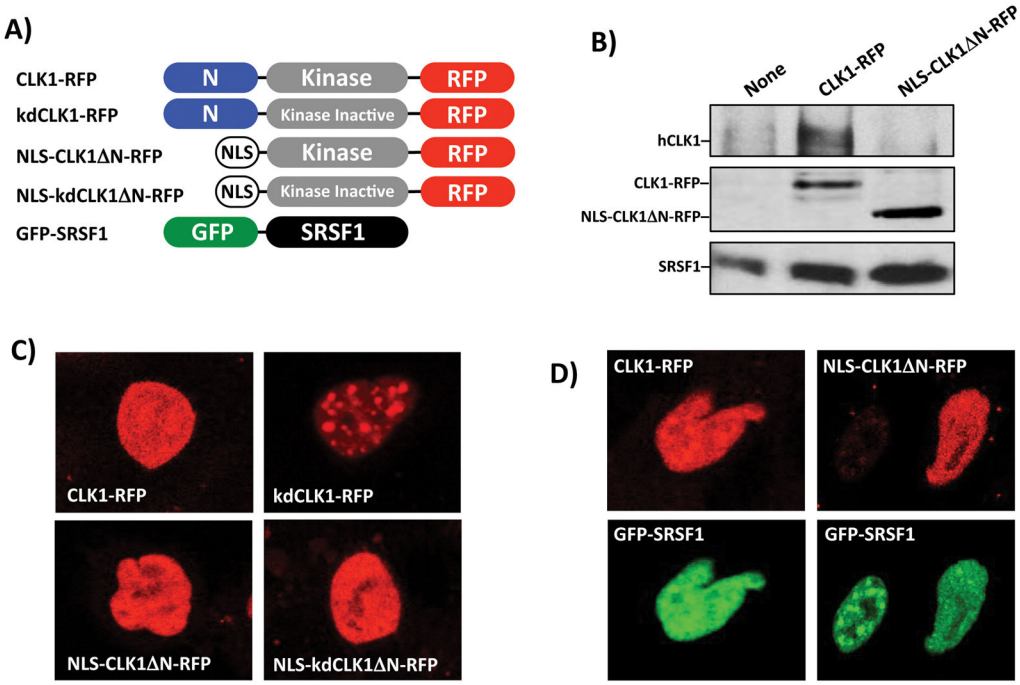

Figure 5. N-terminus affects the subcellular localization of hCLK1 and SRSF1

A) RFP- and GFP-tagged forms of hCLK1 and SRSF1. B) hCLK1 self-associates in HeLa cells. Nuclear lysates of HeLa cells expressing CLK1-RFP are immunoprecipitated with an anti-RFP antibody and probed with a monoclonal antibody for endogenous hCLK1. C) Live-cell confocal imaging of several RFP-tagged forms of hCLK1. D) Live-cell confocal imaging of both CLK1-RFP and GFP-SRSF1 and NLS-CLK1ΔN1-RFP and GFP-SRSF1.

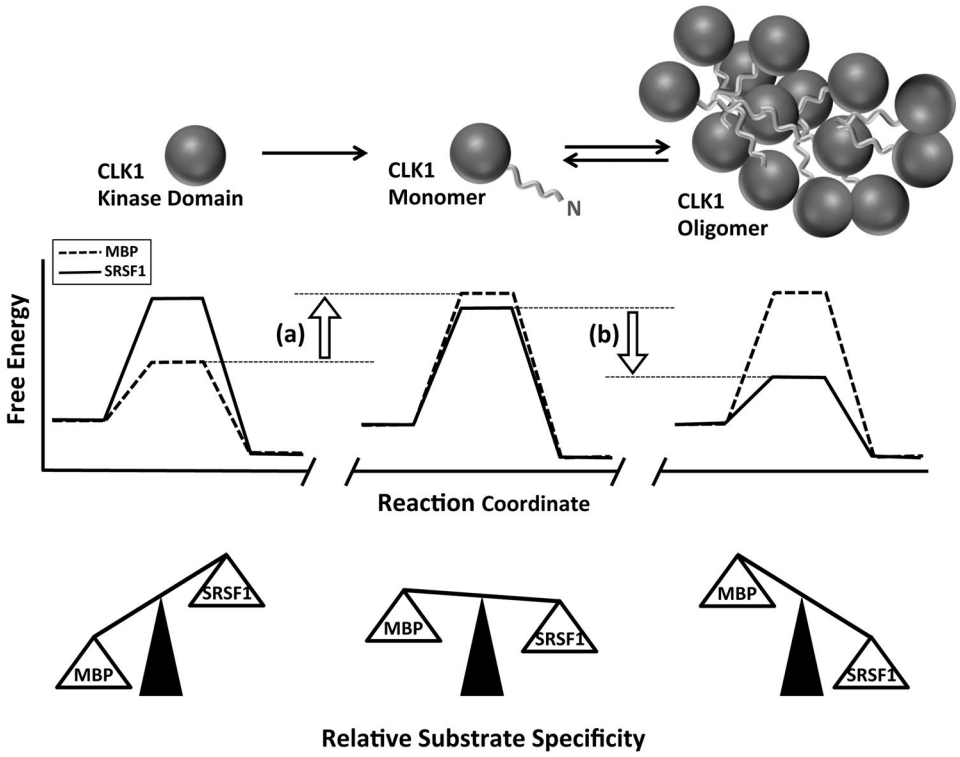

Figure 6. Oligomerization Directs hCLK1 Toward Physiological Substrates

Reaction coordinate diagrams show relative changes in energy barriers for the phosphorylation of SRSF1 and MBP by the hCLK1 kinase domain, monomeric hCLK1 and oligomeric hCLK1. Addition of the N-terminus to the kinase domain raises the energy barrier for MBP phosphorylation (a) whereas oligomerization of the full-length kinase lowers the energy barrier for SRSF1 (b).