Discovery and characterization of autoinhibition in the *S. pombe* decapping complex

By

David R. Paquette

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

In

Biophysics

i

ACKNOWLEDGEMENTS

ABSTRACT

The decapping complex is a dynamic enzyme that is involved in dynamic processes of development and environment stress responses. *In vivo,* and to some extent *in vitro*, decapping has been shown to be modulated by numerous cellular cofactors; including Dhh1 and Pat1:Lsm1-7, which help define the timing of decapping. Additional coactivating sequences contained in lineage specific proteins allow for plasticity in recruiting the complex to novel pathways. Recently, great progress has been made in describing the structured, core domains of the various 5' decay factors and the decapping structure has been solved in multiple forms that represent various conformations along its enzymatic trajectory; including coactivator bound. Much less is known about the intrinsically disorder regions (IDRs) of these proteins. *In vivo* results have implicated that there could be a negative regulatory element in the decapping complex. Following this insight, I have discovered and subsequently characterized an autoinhibitory region of the fungal Dcp2 IDR. Additionally, this autoinhibitory complex reveals a potentially larger role for Edc3 in cytoplasmic decay that had previously been unappreciated. This autoinhibition provides new possibilities for our understanding of mRNA decay and this system provides the opportunity to resolve differences between *in vivo* and *in vitro* decapping results. Finally, our biochemical and structural insights into the core Dcp1:Dcp2 activation have resulted in an enzyme that may prove useful for RNA cloning methods.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

INTRODUCTION

Cells are dynamic in nature and need to be able to adapt rapidly to changes in their environment; be they stressors, cell division or differentiation. Cells have evolved to be able to change their genetic program through several often-redundant pathways. This can be accomplished at the DNA level by changes in nucleosomes that can lead to alterations in the accessibility of the transcription machinery (Narlikar et al. 2002; Tsompana and Buck 2014; Rudnizky et al. 2017); at the RNA level by shuttling transcripts in and out of processing bodies (Decker and Parker 2012) or through 5' or 3' RNA decay (SGrudzien-Nogalska and Kiledjian 2017) or other specialized pathways (Lykke-Andersen and Jensen 2015; Cho et al. 2015); or at the protein level by alterations in localization (Bauer et al. 2015) or turnover by the proteasomal machinery (Ravid and Hochstrasser 2008). Cellular machinery and in particular proteins are also dynamic in nature. Enzymes often exhibit millisecond to microsecond conformational fluctuations during catalytic cycles (Henzler-Wildman and Kern 2007) and these dynamics are often influenced by interactions with other cellular factors: protein cofactors, ligands, etc. The 5' bulk decay pathway, where the decapping complex (Dcp1:Dcp2) plays a major role is nicely at this intersection of a dynamic cellular pathway with dynamic proteins.

5'-3' decay plays a role during the dynamic multi-cellular process of development, and this role appears to be somewhat conserved. In *C. elegans*, RNAi knockdown of the cytoplasmic exoribonuclease, Xrn1, results in a significant ventral closure defect (Newbury and Woollard 2004). In *D. melanogaster*, suppression of Xrn1 results in maldevelopment of the testes among other morphological defects such as a cleft thorax (Nagarajan et al. 2013). Finally in humans, mutations in DcpS (the cap scavenging enzyme) and Edc3 have been implicated in neurological development defects (Ahmed et al. 2014). The Decapping protein 2 (Dcp2) is abundantly expressed in both male and female reproductive organs, suggesting that 5' - 3' decay could be important at some stage of reproductive development. Although more work with

2

differentiated patient stem-cells would have to be performed to more carefully test of the decapping activity is important for these maldevelopment.

Since 5'-3' decay is an important cellular process, in *S. pombe* decapping appears to be essential, organisms have evolved a series of additional cofactors that can serve to alter the decapping complex's activity or its recruitment to specific RNA substrates (Nissan et al. 2010; Jonas and Izaurralde 2013). In general decapping dependent cytoplasmic bulk decay, Dhh1 and Pat1:Lsm1-7 are critical regulators of this process (Coller and Parker 2005; Chowdhury and Tharun 2008). Lesions in either results in the accumulation of deadenylated, capped transcripts consistent with a defect in decapping. Additionally, Dhh1 is critical for maintaining the polarity of codon-optimality dependent mRNA half-lives (Radhakrishnan et al. 2016). Outside of bulk transcript decay, Dcp2 activity can be controlled by its interaction with specialized cofactors. Edc3 in *S. cerevisiae* regulates the stability of only two transcripts, whereas it is unknown how many transcripts it regulates in *S. pombe*. In *D. melanogaster*, Edc3 recruits the decapping complex to the miRNA mediated gene silencing pathway (Eulalio et al. 2007). There also appear to be Edc1-like decapping activators in most eukaryotes where the activator peptide sequence is conserved, but the remainder of the protein is species/lineage specific. This plasticity allows for lineage specific recruitment of the decapping complex (Jonas and Izaurralde 2013). Finally, there are a variety of boutique decay pathways spanning: Non-sense mediated decay, glucocorticoid receptor mediated mRNA decay (GMD), and Pumilio and FBF (PUF) dependent decay. The multitude of pathways makes for a very complex and dynamic process.

Our lab has focused primarily on deciphering how the decapping holoenzyme (Dcp1:Dcp2), in particular its structured catalytic core is regulated. We have shown that the enzyme is dynamic in the absence of substrate and that these dynamics can be modulated by mutations to residues at the interface between the regulatory domain (RD) and the catalytic domain as well as by co-activators. Previous lab members have shown that the RD, Dcp1, Edc1-like activators and W43 (a cap binding residue) all can have

large effects presumably by manipulating the conformations and free-energies of the transitions between them to alter the activity of the enzyme. In the particular case of W43, mutation to alanine was shown to quench the dynamics of the enzyme in solution in the absence of any substrate (Floor et al. 2012). This W43A mutation also prevented Dcp1 or Edc1-like coactivators from stimulating the holoenzyme. Additionally, several published structures of the active and near active form have provided even greater insight into the atomic resolution of how these factors work.

While the structured core of the decapping enzyme is well characterized biophysically, biochemically and structurally, there is a general lack of understanding of the intrinsically disordered regions of a vast number of proteins involved in RNA decay. In our particular system, we know very little of the functional role of the *S. pombe* Dcp2 C-terminal extension despite there being regions of conservation amongst fungi.   Before I began to focus heavily on characterizing these regions of conservation it was only really appreciated that this region of disorder served as a multivalent scaffold for conserved cofactors such as Edc3. There was no defined functional role in terms of altering the enzymatic activity of the decapping complex. Most biochemical studies have removed the C-terminal extension due to its difficulty in purifying and because of the small genetic effect when removed from the endogenous gene under lab controlled conditions.

In the following sections, I will detail the major findings of my thesis research: (1) I will detail the discovery and characterization of a biochemically pure C-terminally extended *S. pombe* decapping complex. This initial discovery and characterization of this more relevant decapping complex has recapitulated some *in vivo* findings of coactivation by Edc3, has shed light onto the nature of the Dcp2 dynamics we have previously measured in solution, and is beginning to inform upon us some longstanding *in vivo* findings in the field that previous *in vitro* decapping systems have not been able to reasonably explain. (2) I will discuss the design and benchmarking of a constitutively activated decapping complex, the super decapping enzyme, that was originally constructed for structural studies, but which we now

detail its use as a replacement for Tobacco Acid Pyrophosphatase (TAP) in 5' end sequencing methods. (3)

I will discuss the results of using an insect cell expression to purify full length Dcp2 and some preliminary

S. pombe genetics to better define the functional role of 5' mediated decay in fission yeast. (4) Finally, I

will provide some speculation on the future of our understanding of the decapping complex.

CHAPTER 1

The mRNA decapping enzyme complex is autoinhibited and activated

by Edc3

David R. Paquette[1,3], Ryan W Tibble[2,3], and John D. Gross[1,2,3]

Author affiliations:

[1] Integrative Program in Quantitative Biology, Biophysics division, University of California, San Francisco, California, 94158

[2] Program in Chemistry and Chemical Biology, University of California, San Francisco, California, 94158

[3] Department of Pharmaceutical Chemistry, University of California, San Francisco, California, 94158

ABSTRACT

5' mediated cytoplasmic RNA decay is a critically conserved cellular process in eukaryotes. While the structured globular domains of the nucleases and cofactors in this pathway have been determined and functionally characterized, our understanding of the abundant intrinsically disordered regions (IDRs) is lacking. Here we reconstitute and describe for the first time a C-terminally extended Dcp2 in complex with Dcp1, which is autoinhibited. We identify two inhibitory motifs (IM1 and IM2) that when removed restore the activity of the Dcp1:Dcp2 complex. The addition of Edc3, which binds additional conserved regions of the IDR, fully alleviates the autoinhibition, stimulating decapping ~40—fold. This robust activation is the most relevant biochemical reconstitution to date and furthers our understanding of Edc3. Using the available crystal structure of Dcp1:Dcp2 to guide mutagenesis, we show that a closed cap-occluded conformation, is consistent with an autoinhibited state. Mutation of a conserved tyrosine alleviated the decapping inhibition, altered the conformational dynamics, and might increase m7G-cap affinity. Our data reveals a conserved cap-occluded state that can be modulated by conserved cofactors and provides a framework for understanding the control of decapping in higher eukaryotes.

INTRODUCTION

Eukaryotic 5'-3' mRNA decay is preceded and permitted by removal of the m7GpppN (N any nucleotide) cap structure and is a critical, conserved cellular process from yeast (Dunckley and Parker 1999) to humans (Wang et al. 2002; SGrudzien-Nogalska and Kiledjian 2017). The maintenance of the m7G cap in eukaryotes is required for a variety of cellular processes including: splicing (Konarska et al. 1984; Izaurralde et al. 1994), nuclear export (Visa et al. 1996), canonical translation and transcript stability (Moore 2005; Topisirovic et al. 2011); defective caps are removed by quality control pathways (Xiang et al. 2009; Chang et al. 2012); and recent evidence suggest caps containing m6Am as the first-transcribed nucleotide are protected from decapping and 5'-3' decay (Mauer et al. 2016). Critically, mRNA decay is important during development (Newbury and Woollard 2004). Additionally, the cap structure and polyA tail differentiate the mRNA from other cellular RNAs; a feature viruses try to coopt to protect and translate their messages (Molleston and Cherry 2017). Since ablation of the cap structure removes the message from the actively translating pool and marks it for turnover, the 5' decapping pathway has evolved to rely on a host of cofactors (Arribas-Layton et al. 2013). A continual outstanding question remains in how the activity of the decapping enzymes are controlled.

There are a variety of decapping enzymes in eukaryotic cells (Song et al. 2010). The decapping enzyme Dcp2 has been implicated in bulk 5'-3' decay, mRNA quality control and regulated mRNA decay pathways (Schoenberg and Maquat 2012), such as the decay of miRNA targets and glucocorticoid receptor mediated decay (Cho et al. 2015). Dcp2 is a bi-lobed enzyme made-up of a regulatory and NUDIX containing catalytic domain, and Dcp1, an EVH1-like scaffold, which recruits cofactors through an aromatic cleft that recognizes a short proline-rich motif (She et al. 2008). Additionally, fungal Dcp2 contains a long C-terminal extension. This extension has been shown to recruit several known regulators of decay (He and Jacobson 2015). The most well characterized of these interactions is between short helical leucine-

rich motifs (HLMs) and enhancer of decapping 3 (Edc3). In budding yeast, deletion of Edc3 causes upregulation of the YRA1 pre-mRNA and the mRNA encoding the ribosomal protein RPS28B. It also functions in assembly of P-bodies. In yeast Edc3 plays a role in localizing Dcp1/2 to processing-bodies (Decker et al. 2007; Fromm et al. 2012) and *in vitro* is sufficient to facilitate liquid-liquid demixing (Fromm et al. 2014). Quite possibly Edc3 may act as a general enhancer of decapping in fission yeast, since deletions of Edc3 resulted in decreased decapping of an abundant mRNA and an increase in the number of processing bodies (Wang et al. 2013). In Drosophila , Edc3 is involved in the decay of some miRNA targets, whereas in humans, a single mutation in its LSm domain (HLM binding) is highly correlated with defects in neuronal development (Ahmed et al. 2014). These interactions facilitated by the intrinsically disordered regions (IDRs) are highly plastic, evolve rapidly, and facilitate the rewiring of mRNA decay pathways amongst species (Jonas and Izaurralde 2013).

Recently, a region in the fungal Dcp2 specific extension was described (He and Jacobson 2015), which imparted a negative effect on a subset of Edc3 specific transcripts in yeast depleted for Edc3. Removal of the *cerevisiae* conserved sequence restored transcript levels, bypassing the Edc3 requirement. However, it is unclear from these studies if the region is a direct inhibitor of decapping or if additional factors are involved. How inhibitors of decapping modulate the complex conformational landscape of Dcp2 is unknown. Furthermore, this specific sequence is not found in the distally related fission yeast. Here we provide the first direct evidence that a C-terminal extended Dcp1:Dcp2 complex from S. *pombe* is autoinhibited. Additionally, a proline-rich region contains two sub-regions of conservation that when removed fully restore enzymatic activity. Third, we demonstrate that the addition of Edc3 alleviates the inhibitory role of the disordered region. Fourth, we identify a critical amino acid, Y220, which when mutated quenches ms/μs dynamics of Dcp2, restores activity of the inhibited complex, and bypasses the Edc3 mediated alleviation of the inhibition. Taking all of the available data into account, we propose a model for the autoinhibition of the decapping complex through the stabilization of a cap-occluded

conformation that has appeared in numerous structural studies and was suggested to be an inactive form of the enzyme.

RESULTS

The disordered c-terminal extension of Dcp2 is autoinhibitory

The C-terminus of pombe Dcp2 is predicted to be highly disordered (**Fig 1A**). We were interested in determining if this largely disordered extension had a functional role. Previously, a region containing proline-rich sequences (PRS) was shown to interact with the aromatic quartet of Dcp1 (She et al. 2008; Borja et al. 2011; Lai et al. 2012; Valkov et al. 2016; Mugridge et al. 2016), a binding site for decapping coactivators. However, unlike Edc1,2/PNRC2 there was no apparent effect on decapping rates when the sequence was added in excess *in trans*. To get at this question of a functional role, we coexpressed and purified Dcp1 and a C-terminally extended Dcp2. The purified complex was compared side-by-side with the catalytic core in an *in vitro* decapping assay. It was apparent from the raw TLC plate that the enzymatic activity was reduced (**Fig 1B**). When the data was plotted and the rates determined, we found that the extended complex was consistently 10-fold slower than the structured core decapping complex alone (**Fig 1C,D**). This effect was not dependent enzyme preparation, and the larger decapping complex containing the inhibitory region purified as a well-resolved, homogeneous peak on gel-filtration (**Fig S2**)

The disordered extension contains a c-terminal regulatory region (CRR) composed of two inhibitory motifs

Having determined that Dcp1:Dcp2(1-504) was less active than the functional core, we wanted to know which region encoded for this effect. We narrowed our search window by querying the

10

conservation amongst the most closely related fission yeast. The rationale being that motifs in highly disordered primary sequence evolve rapidly due to less constraint, i.e. lack of structural confinement, and an ability to swap from one protein to another (Davey et al. 2015; Latysheva et al. 2015). Using this approach, we identified a critically conserved stretch of amino acids and the previously characterized PRS (**Fig 2A**). While previous studies have failed to show any inhibitory effect for this PRS, we sought to see if in the context of our inhibited complex ablations of these conserved regions would contribute to the decrease in decapping activity. We purified to homogeneity Dcp1:Dcp2 complexes where either the PRS or region 2 alone or in tandem were deleted (**Fig S2B**). Unlike in the trans addition experiments, the PRS contributes partially to the autoinhibition when present in cis. Likewise, excision of region 2 restored activity to a similar extent. More strikingly, the combined removal of these regions fully restored the activity to that of the catalytic core (**Fig 2B,C**). We therefore have named these regions inhibitory motif 1 (IM1) and inhibitory motif 2 (IM2) respectively. The region comprising IM1 and IM2 we now consider to be a C-terminal regulatory region or CRR.

Alleviation of autoinhibition by the conserved Edc3

With this new CRR mapped, we wondered if any of the known regulators relieved this inhibition. Our longer Dcp2 construct contains at least three conserved HLMs that are known binding sites for Edc3 (Fromm et al. 2014), so it seemed plausible that Edc3 might alleviate autoinhibition. When Dcp1:Dcp2 (1-504) was incubated with excess Edc3, the activity was completely restored to levels beyond the catalytic core alone (**Fig 3A,B**). Furthermore, this activity was dependent on the known LSm domain, since the YjeF N domain alone did not alter the decapping activity (**Fig 3B**). Interestingly, addition of the LSm domain alone resulted in an increase in activity that was on par with the core decapping complex (**Fig 3A,D**). These results suggest that binding the HLMs with the LSm domains was sufficient to block the autoinhibition via

the CRR. Full-length Edc3 provides an additional level of activation of Dcp1:Dcp2 (**Fig 3B**). Prior studies of in *S. cerevisiae* Dcp1:Dcp2 revealed the core structure domain and an additional supported modest levels of Edc3 activation of decapping Edc3 (Harigaya et al. 2010), which was not due to autoinhibition. Accordingly we compared fold activation of decapping of Dcp1:Dcp2 (1-266) (**Fig 3B**) with the substantially longer autoinhibited construct. The LSm domain does not provide any substantial increase in activity of Dcp1:Dcp2, and only an approximate 5-fold increase upon addition of Edc3 (**Fig 3D**). This implicates that only in the context of the longer auto-inhibited Dcp1:Dcp2 complex do we see a robust activation by Edc3 with the majority of the change coming from the recognition of Dcp2 by the HLM binding LSms, with an additional 4 to 5 -fold coming from the full Edc3.

Cap-occluded Dcp2 conformation is a vestige of an inactive state

We and others have suggested that the ATP bound, closed Dcp2 structure could resemble an inactive conformation of the holoenzyme Dcp1:Dcp2 (Floor et al. 2010; Fromm et al. 2012). Indeed, a recent crystal structure of Dcp1:Dcp2 and coactivators with substrate analog and product indicate an extensive remodeling from a cap-occluded state to one that is accessible for interactions with cap (She et al. 2008; Valkov et al. 2016; Mugridge et al. 2016). However, the decapping activity of the structured core domains of Dcp1:Dcp2 that were previously studied by crystallography are only modestly enhanced upon addition of the LSm domain (~1.5-fold, **Fig 3D**). Nevertheless, we hypothesized that the autoinhibited state might mimic a cap-occluded conformation and that the regions of the CRR may consolidate this state. The Dcp1:Dcp2 cap-occluded structure is stabilized by several conserved residues, including: W43, D47 and Y220 (**Fig S4A**). In the m^{7}GDP bound state, the Y220 was no longer close to the active site (**Fig 4A**). Upon mutation of Y220 to glycine, the activity of Dcp1:Dcp2 (1-504) was restored to within ~2-fold of the catalytic core (**Fig 4B**). These data suggest the cap occluded conformations of Dcp1:Dcp2 might

represent the autoinhibited form of the enzyme which is further stabilized by the CRR. Consistent with

this interpretation, the Y220G mutation nearly bypassed the requirement for Edc3 activation; the change

in activity being more similar to that of Dcp1:Dcp2(1-266) than Dcp1:Dcp2(1-504) (**Fig 4C**).

Mutation of Y220 quenches dynamics and may increase cap affinity

Previously, it had been noted that Dcp2 experiences ms/µs concerted dynamics across the RD and

CD. W43 was identified as a "gatekeeper" residue (Floor et al. 2012). Since Y220 makes interactions with

W43, we thought that the Y220 mutation would also perturb these dynamics resulting in the

reappearance of hallmark cross-peaks (**Fig S6**). Upon mutation of Y220 to glycine in Dcp2(1-243), the

expected cross-peaks returned from across both the RD and CD along with W43, which is undetectable in

the WT spectrum (**Fig 5A**). A second prediction of destabilizing the cap-occluded conformation is higher

affinity for m7GDP. Monitoring the W43 indole cross-peak, we were able measure K_D of ~3mM, upon

titrating m7GDP, compared to ~ 10mM for the individual domains (**Fig 5B**). While we were able to quantify

cap binding to Y220G, we could not reliably measure m^7GDP binding by CSP in the WT protein. This was

mainly in part due to the ms/µs dynamics, general broadening, and peaks exhibiting multiple exchange

regimes. Since we were not able to measure this affinity directly, we reasoned that increased cap affinity

compared to WT would result in a concomitant increase in the enzymatic rate. In agreement, we

consistently measured a 1.5 fold enhancement in the rate for the Y220G compared to WT (**Fig 5C**)

DISCUSSION

mRNA decay factors have been extensively characterized biochemically, genetically, and

structurally (Garneau et al. 2007; Schoenberg and Maquat 2012). In the past decade or so, several studies

have provided insight into the mechanisms of how Dcp1 and its co-activators (Edc1-like) are able to activate the decapping holoenzyme (Borja et al. 2011; Lai et al. 2012; Valkov et al. 2016; Mugridge et al. 2016; Wurm et al. 2017). Furthermore, a plethora of structures of a majority of the globular domains of the mRNA decay proteins are available and have provided a relatively clear picture of how these domains function (Jonas and Izaurralde 2013). While much progress has been made on this front, much less is known in regard to the intrinsically disordered regions (IDRs) that are replete in these proteins. Here we have determined a CRR of the IDR in *pombe* Dcp2 is autoinhibitory in the context of the Dcp1:Dcp2 holoenzyme. Consistent with the IDR being of functional importance, removal of two fission yeast conserved motifs (IM1 and IM2), resulted in a complete restoration of the enzymatic activity. Importantly, we have shown that the conserved Edc3 is able to alleviate this autoinhibition and that this activation is 25-fold greater than what has previously been reported (Fromm et al. 2012). Finally, we were able to gain insight into the structure of the inhibited state based upon multiple available cap-occluded structures (**Fig S4**), and were able to identify Dcp2 Y220 as a key residue in stabilizing this cap-occluded conformation.

IDRs of proteins are the dark matter of protein sequence space. While the majority of protein folds have been characterized, much less is known about the disordered regions between domains and extending from them. This is despite IDRs making up a sizeable portion of the proteome (Tompa 2003). These regions are an active area of research, since IDRs are plentiful in RNA decay proteins despite diverging considerably in length and composition in eukaryotes, while the globular domains are highly conserved (Jonas and Izaurralde 2013). With the high plasticity and rapid evolution of the IDRs, they provide an opportunity for species specific rewiring resulting in novel functions. The interaction of Edc3 with the IDR of Dcp2 is one of the most well characterized SLiMs mediated complex formation in RNA decay. In seminal work, He and Jacobson demonstrated that in addition to providing a PPI, the IDR of Dcp2 could negatively regulate decay in an Edc3 dependent manner (He and Jacobson 2015). However, it was not determined whether this was a direct affect or if it required additional factors. In previous *in*

vitro studies, a large section of the Dcp2 IDR was removed due to expression and stability issues (Fromm et al. 2014; Schütz et al. 2017). We reconstituted and quantified the direct inhibition of Dcp2 by this IDR in vitro using recombinant purified proteins and further have shown that autoinhibition can be modulated in an Edc3 dependent manner. It will be interesting to see if a similar negative regulation is present in higher eukaryotes, where the IDR containing the Edc3 protein interaction motifs have been transferred to Dcp1 (Fromm et al. 2012).

Assembling the available evidence from the literature and from this study, we propose a model for autoinhibition within the decapping holoenzyme (**Fig 6**). The central aspect of the autoinhibitory conformation is the cap-occluded structure, where Y220 is making edge interactions with the critically important W43 and D47 in a conformation that is incompatible with cap binding (Floor et al. 2010). In this conformation, the CRR likely makes interactions with the catalytic core of the Dcp1:Dcp2 complex; thereby, stabilizing the cap-occluded conformation. Consistent with this model, the region of IM1 has been reported to bind the hydrophobic cleft of Dcp1 (Wurm et al. 2016). Since binding of IM1 in *trans* did not affect the activity of Dcp1:Dcp2, it would make sense that some region of the IDR might contact a domain of Dcp2. We favor that IM2 binds the CD, while the upstream IM1 is engaging Dcp1. The other aspect of this system is in how Edc3 engages and alleviates the inhibition within Dcp1:Dcp2. The most parsimonious mechanism is that in binding to the HLMs in the IDR of Dcp2, Edc3 blocks the CRR from binding to the structured Dcp1:Dcp2 core. Consistent with this mechanism, just the LSm domain was able to restore the activity to catalytic core levels. The full-length Edc3 provided the same effect but with additional activation most likely do to a substrate RNA binding effect, since a dimer of Edc3 has previously been shown to effectively bind RNA (Ling et al. 2008). This is even more apparent if one compares the fold activation of Dcp1:Dcp2 (1-266) with either the LSm domain or the full Edc3. The (1-266) construct is stimulated by 1.3-fold by the LSm domain, whereas full-length Edc3 stimulates ~4-fold. Arguably, the stimulation to decapping by Edc3 observed in the structured core domains of Dcp1:Dcp2 containing a

single HLM is driven by the additional RNA binding surface provided by Edc3. In contrast in the longer autoinhibited constructs of Dcp2 containing the CRR, Edc3 provides both an increase in activity of Dcp1:Dcp2 by blocking inhibition (~10-fold) and by increased RNA binding capacity (~4-fold). With this larger fold activation effect for Edc3 and C-terminally extended Dcp2, a challenge for the future is to determine the identities of the transcripts influenced by Edc3.

Previously, our lab had concluded that Dcp2 undergoes a collective fast-exchange between open and closed states in the absence of substrate (Floor et al. 2012). We interpreted this as Dcp2 dynamically sampling an open, inactive and closed, active states with W43 playing a critical gatekeeper role. Here we have identified a residue, Y220, in the CD lobe of Dcp2, which when mutated affected cross-peaks similarly (**Fig S6**). With a high similarity in the cross-peaks that reappeared and shifted, we expect that Y220G also quenches these ms/µs dynamics. Y220 mutation appears to destabilize the cap-occluded structure. Consistent with this view, Y220G increased the catalytic rate of the Dcp1:Dcp2Dcp2 (1-243) catalytic core by 1.5-fold over WT. Also, consistent with the inhibited state being related to this cap-occluded state, Y220G restored activity to the longer autoinhibited Dcp2. Taking these enzymatic effects and the NMR data into account, we suggest that the dynamics Dcp2 exhibits correspond to a rapid sampling between a cap-accessible and a cap-occluded, inhibited state. It would be of interest to understand how this cap-occluded to cap-accessible transition is perturbed by Dcp1 and other decapping cofactors.

Edc1-like activators bind a conserved cleft in Dcp1 and stimulate the catalytic step of the Dcp1:Dcp2 holoenzyme (Borja et al. 2011). Since the IM1 of the C-terminal regulatory region interacts with the same Dcp1 surface, it will be of interest to see how these activators are impacted by the autoinhibited conformation. Can the Edc1-like activators alleviate the autoinhibition, and if so is the concentration of the activator required for robust activation shifted? Or is it possible that the Edc1-like activators are not functional in the absence of Edc3? Furthermore, in *cerevisiae* additional mRNA decay factors have been shown to bind the IDR of Dcp2 ((He and Jacobson 2015). Do any of these additional

factors bind to the *pombe* IDR, and if so, do they also alleviate inhibition by Edc3. Lastly, how do higher eukaryotes utilize the cap-occluded state. Future work will be required to mechanistically understand how the inhibited state is controlled and nature of the contacts, which stabilize the autoinhibited conformation.

Materials & Methods

Protein expression and purification

A pRSF containing polycistronic His-Gb1-tev-Dcp1:Dcp2(1-504)-strepII was used to coexpress the c-terminally extended Dcp1:Dcp2 complexes. Both the Dcp1 and Dcp2 sequences were codon-optimized for *Escherichia coli* from Integrated DNA Technologies and were cloned into a pRSF vector using Gibson assembly. The Dcp2 sequence contained a ribosome binding site (rbs) upstream and the Dcp1 was cloned behind the endogenous T7 promotor and rbs within the vector. The His-Gb1-tev-Dcp1:Dcp2(1-504)-strepII were expressed in *E. coli* BL21(DE3) (New England Biolabs) grown in LB medium. Cells were grown at 37 °C until they reached an OD_{600} = 0.6-0.8, when they were transferred to 4 °C for 30 minutes before induction by addition of 0.5mM IPTG (final concentration). Cells were induced for 16-18 hours at 18 °C. Cells were harvested at 5,000*g*, lysed by sonication (50% duty cycle, 3x1min), and clarified at 16,000g in lysis buffer (buffer composition goes here). The protein complex was purified in a two-step affinity purification: Ni-NTA agarose affinity column followed by strep-tactin high-capacity superflow and elution with 5mM d-desthiobiotin. The His-Gb1 tag was removed by addition of TEV protease overnight at 4 °C. The complex was further purified by size-exclusion chromatography on a GE Superdex 200 16/60 column in storage buffer (50mM HEPES, 100mM NaCl, 5% glycerol, 5mM DTT pH 7.5). The purified complex was concentrated, 20% v/v (final) glycerol was added, and then flash frozen in LN2 for kinetics studies.

Dcp1:Dcp2(1-504) internal deletion constructs (IM1, IM2 and IM1/IM2) were generated by whole-plasmid

pcr with 5'Phosphorylated primers (source). Dcp1:Dcp2 (1-504) Y220G was generated using whole-

plasmid pcr with mutagenic divergent primers. A His-TRX-tev-Edc3 containing plasmid was generated by

Gibson cloning *S. pombe* Edc3 cDNA into a pET30b plasmid which already contained a His-TRX-tev coding

sequence. The LSm and YjeF N domain containing plasmids were created by whole-plasmid pcr with

5'Phosphorylated primers. The Edc3 constructs were purified as described (Fromm et al. 2014) with a

modification to the size-exclusion chromatography; storage buffer was used instead of the phosphate

buffer as described.

Kinetic decapping assays

Single-turnover in vitro decapping assays were carried out as previously described (Jones et al.

2008). Slight alteration to the buffer (brought up volume in storage buffer with additional 20% v/v

glycerol) A ^{32}P-labeled 355-mer (A15 tail) RNA substrate derived from MFA2 was used for all of the

decapping assays. Reactions were initiated by mixture of 30 µL 3× protein solution with 60 µL 1.5× RNA

solution at 4 °C; final Dcp1:Dcp2 concentration was 1.5uM and the final RNA concentration was <100 pM.

For decapping assays containing Edc3; Edc3 was added in 4-fold molar excess and the mixture as

incubated at RT for 20 minutes before transferring to the 4 °C block. Samples were equilibrated for at

least 30 minutes on the 4 °C block before initiating. Time points were quenched by addition to excess

EDTA, TLC was used to separate the RNA from the product m^{7}GDP, and the fraction decapped was

quantified with a GE Typhoon scanner and ImageQuant software. Fraction m^{7}GDP verus time were

plotted and fit to a 1st-order exponential to obtain kobs; in the case of Dcp1:Dcp2(1-504) when the kinetics

were too slow to obtain reliable exponential fits, kobs was obtained from a linear fit of the initial rates by

division of the slope by the empirically derived endpoint.

NMR related

[15]N-labeled *S. pombe* Dcp2 with a N-terminal His-GB1-TEV tag was expressed in *E. coli* grown in M9 minimal media containing ^{15}N-ammonium chloride ([15]NH$_4$Cl) as the sole nitrogen source. Following overnight induction with 1 mM IPTG, cells were lysed by sonication and clarified at 16,000*g*. Dcp2 was purified by incubating clarified lysate with nickel resin followed by elution with 250 mM imidazole. The His-GB1-TEV tag was then removed by digestion with TEV protease and untagged Dcp2 was run over a Superdex 75 size exclusion chromatography column equilibrated with pH 7.0 NMR buffer (21.1 mM NaH$_2$PO$_4$, 28.8 mM Na$_2$PO$_4$, 200 mM NaCl, 100 mM Na$_2$SO$_4$, 5 mM DTT).

All ^{1}H-^{15}N HSQC experiments were performed at 303K on either a Bruker Avance DRX500 or Bruker Avance 800 spectrometer equipped with cryogenic probes. For NMR titrations of m[7]GDP, Dcp2 was exchanged into nucleotide binding buffer (50 mM Hepes pH 7.0, 150 mM NaCl, 2 mM MgCl$_2$, 5 mM DTT) using a BioRad desalting column. Titrations were carried out at 100 µM protein with equimolar addition of m[7]GDP and MgCl$_2$. Chemical shift perturbations were determined using the standard Euclidian equation (1):

$$\sqrt{\left(\delta H_{bound} - \delta H_{free}\right)^2 + \left(0.2\left(\delta N_{bound} - \delta N_{free}\right)\right)^2}$$

where the factor of 0.2 is used as a scaling factor for the nitrogen spectral width. Titration data were fitted to the standard hyperbolic binding equation if the observed perturbation between apo and 20 mM m[7]GDP for any residue was greater than the mean plus one standard deviation. Experimental K$_D$ values are reported as ± standard error.

Figure1: The dcp2 disordered C-terminal extension is autoinhibitory

A.

B.

C.

D.
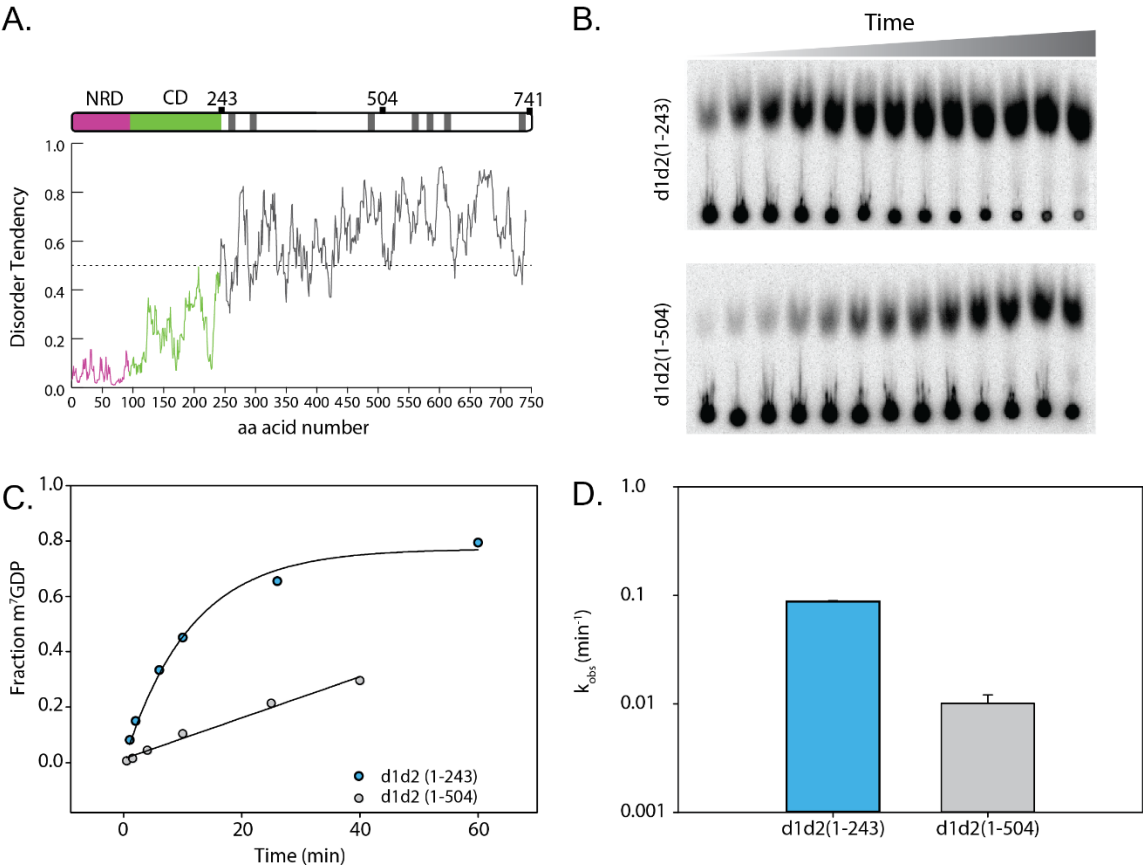
Figure2: Mapping regions in the C-terminal regulatory region that contribute to auto-inhibition
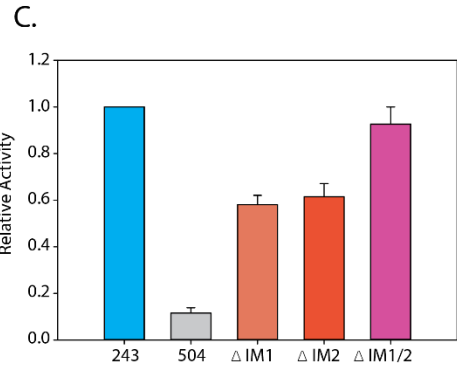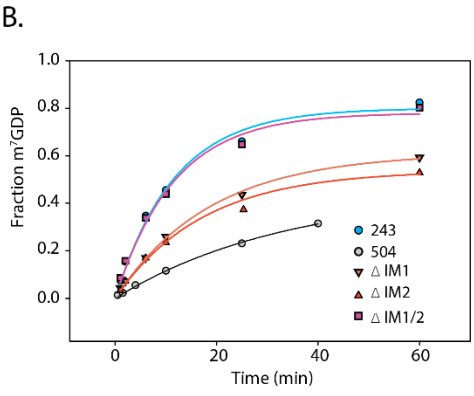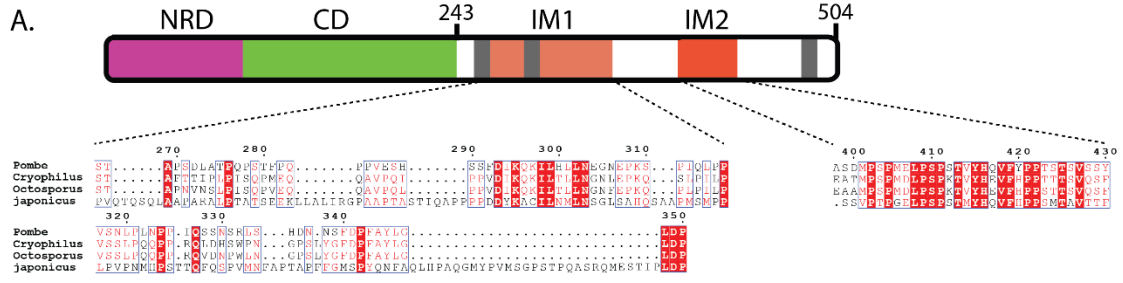
Fig 3: Edc3 alleviates Dcp1:Dcp2 autoinhibiton

Figure4: Y220 stabilizes a cap occluded state and alleviates inhibition

A.
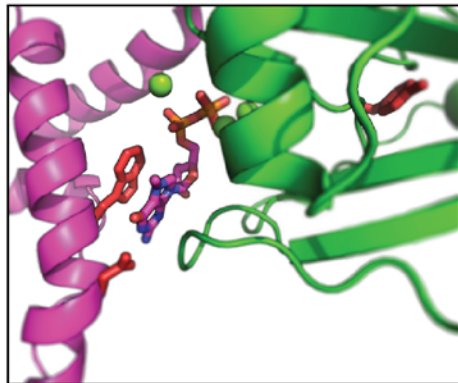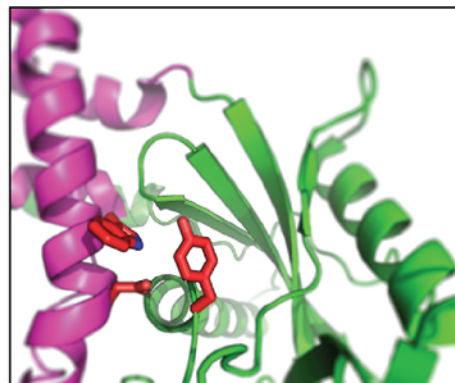
Cap - accessible   Cap - occluded

B.
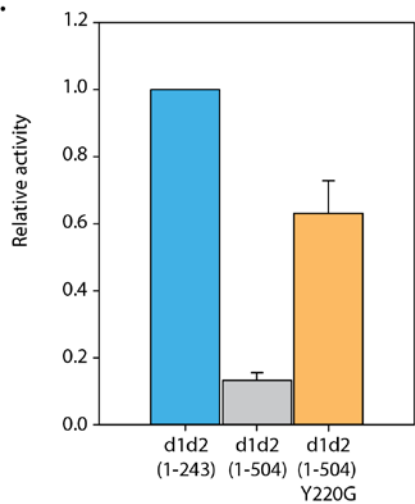
C.
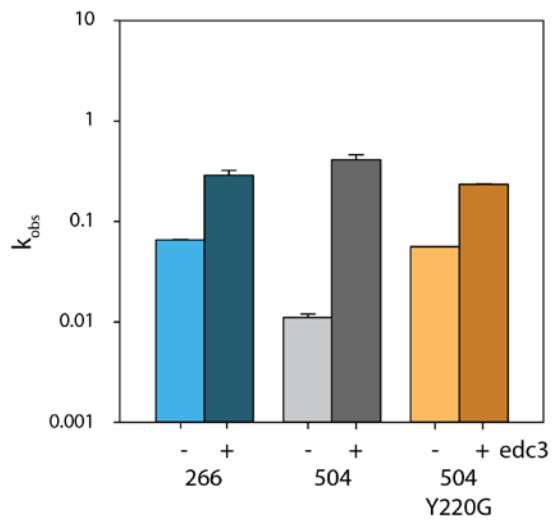
Figure 5: Y220G mutation perturbs the WT HSQC spectrum and may increase cap affinity

A.
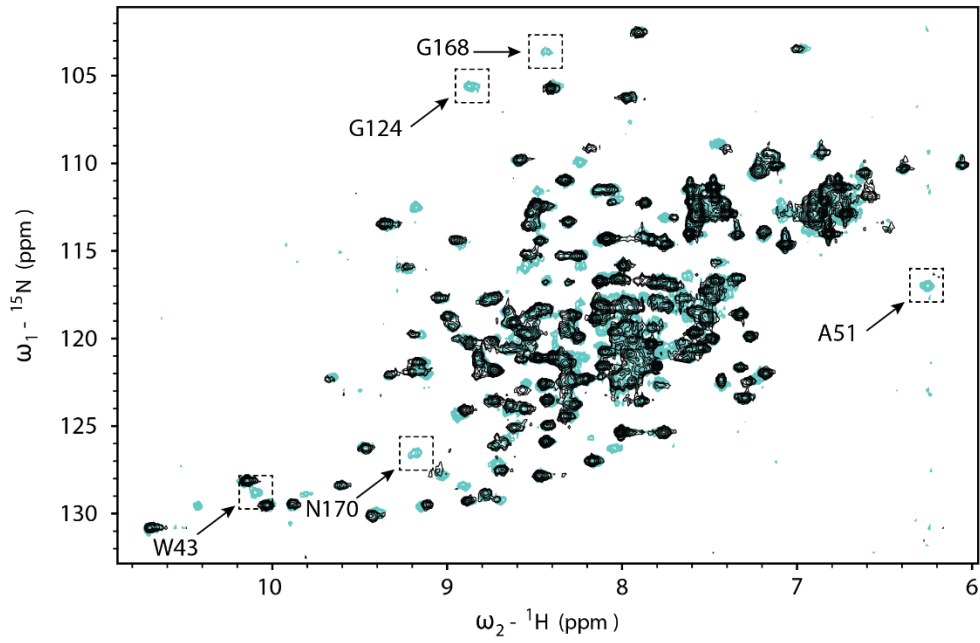
B.

C.

Fig6: Model of dcp1:dcp2 autoinhibiton and edc3 activation

$k_{accessible}$

$k_{occluded}$

k_{chem}

edc3 ✓

edc1-like ?
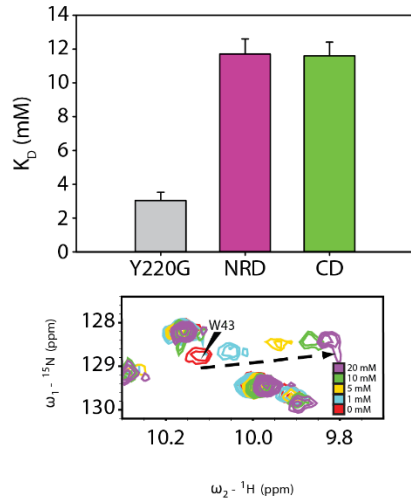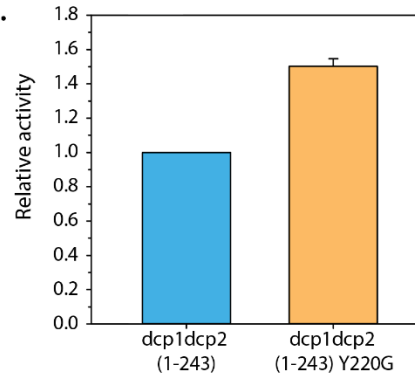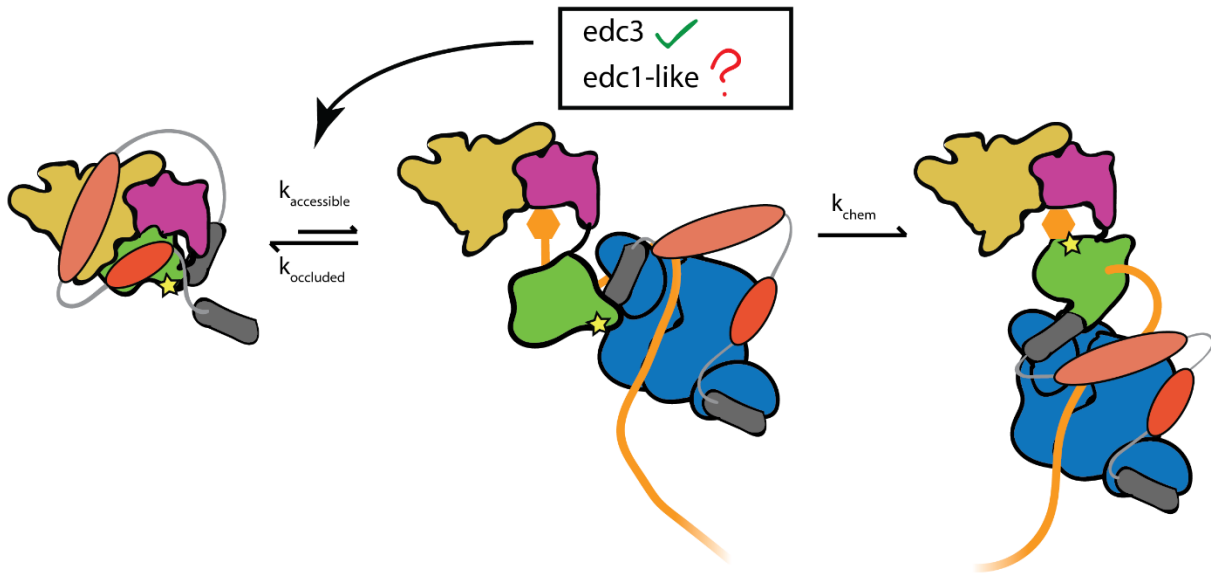
MAIN FIGURE LEGENDS

Fig 1: The Dcp2 c-terminal extension is disordered and autoinhibitory

(A) Block diagram of the domains of *pombe* Dcp2. The magenta box (1-94) comprises the N-terminal regulatory domain (NRD) and the green (95-243) comprises the catalytic domain, which contains the Nudix helix. Grey bars are helical leucine-rich motifs (HLMs). The disorder tendency plot below was calculated using the IUPRED server (Dosztányi et al. 2005). Where regions above the dotted line have a higher propensity for being disordered. (B) Representative raw TLC (thin-layer chromatography) showing the formation of m^7GDP product over time. The lower spots are the RNA origin and the upper are the cleaved cap. (C) Representative plot of fraction m^7GDP product versus time comparing the catalytic core and inhibited C-terminally extended Dcp1:Dcp2. (D) A log-scale plot of the empirically determined rates from (C), where the error bars are the population standard deviation, σ.

Fig 2: The disordered region or Dcp2 contains a c-terminal regulatory region (CRR) consisting of two inhibitory motifs (IMs). (A) Block diagram colored as in Fig1. with IM1 and IM2 regions colored and the sequence conservation (Robert and Gouet 2014) for each motif shown below. IM1 is generally proline rich, while IM2 is absolutely conserved in all fission yeast. (B) Plot with fits for fraction of m^7GDP versus time comparing the activity of Dcp1:Dcp2(1-504) where either IM1, IM2 or both are internally deleted. (C) Bar graph of the relative enzymatic activity of the various Dcp1:Dcp2 complexes compared to the (1-243) catalytic core. Each IM contributes to the inhibitory effect of the CRR. The error bars are the population standard deviation, σ.

Fig 3: Edc3 alleviates the inhibitory effect of the Dcp2 CRR. (A) Block diagram of the Dcp2 and Edc3 used in the subsequent decapping assays. Edc3 consists of an LSm domain that interacts with HLMs and a YjeF N domain that provides an RNA binding surface when dimerized. (B) Decapping activity of Dcp1:Dcp2 (1-504) incubated with excess Edc3, LSm domain or YjeF N domain. (C) Same as B but with Dcp1:Dcp2 (1-266) instead. (D) Comparison of the relative fold activation of Dcp1:Dcp2 (1-504) versus (1-266) with the various Edc3 constructs. The error bars are the population standard deviation, σ.

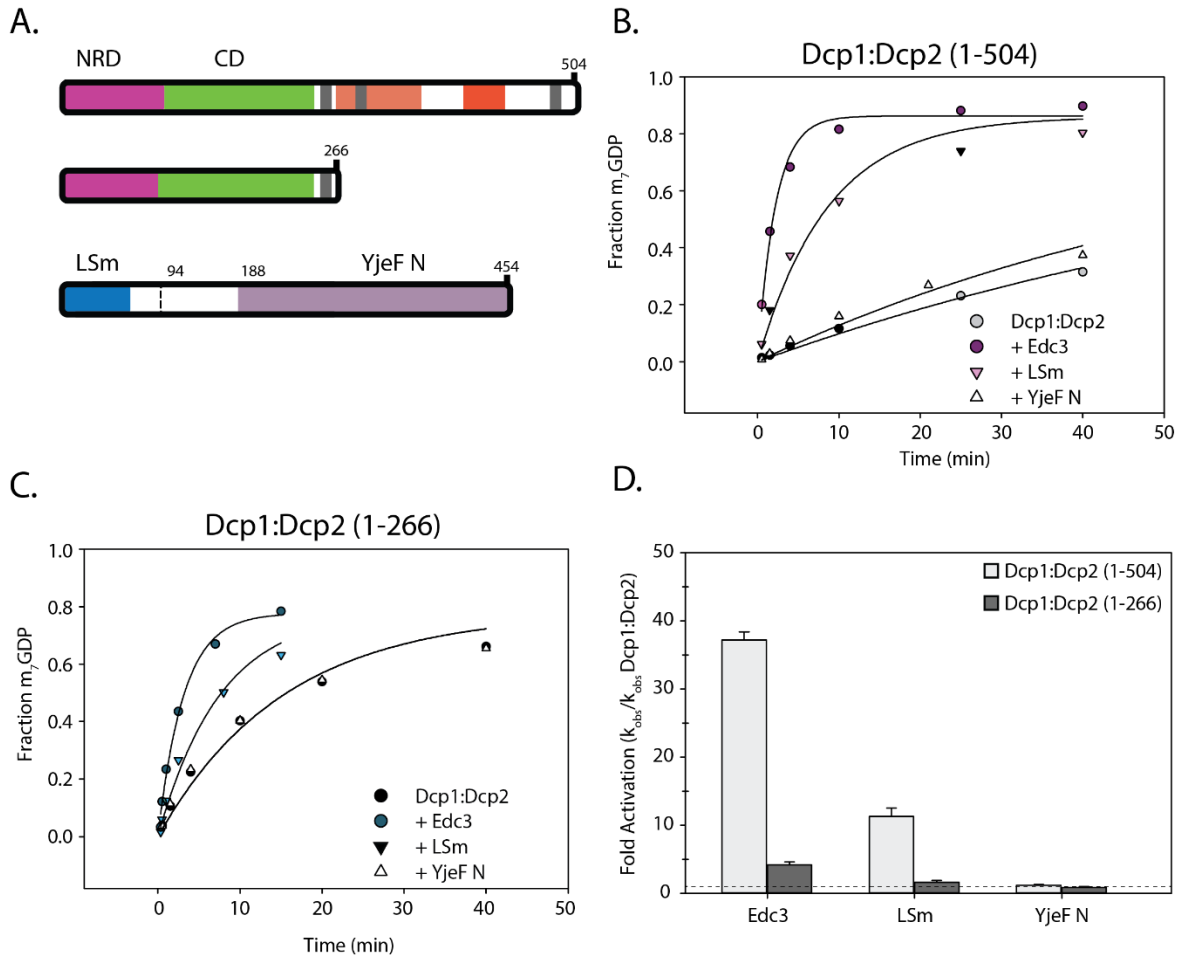**Fig 4:** Y220 stabilizes a cap-occluded state and alleviates inhibition (A) W43 and D47 that coordinate the m^7G cap exist in conformations where they are either accessible or occluded by interaction with the conserved Y220G. (B) Plot of the relative activity of WT or Y220G Dcp1:Dcp2(1-504) compared to the structured catalytic core. (C) Logscale plot of decapping rate of Dcp1:Dcp2 (1-266) [blue], (1-504) [gray], or (1-504) Y220G [yellow] where the darker bar is the rate with excess Edc3.

Fig 5: Y220G mutation perturbs the WT HSQC spectrum similarly to W43A and may increase cap affinity. (A) Shown are the ^{15}N HSQC spectra of WT Dcp2 residues 1-243 (black) and Dcp2 Y220G (cyan). Selected residues with significant changes upon mutation are indicated. (B) Plot showing the K_D of m^7GDP for Dcp2(1-243) Y220G and the isolated domains of WT NRD or CD; below is a the CSP trajectory for the W43 indole upon m^7GDP titration into Y220G. (C) Dcp1:Dcp2 (1-243) Y220G mutation is ~1.5-fold faster than the WT protein in an *in vitro* decapping assay.

Fig 6: Model of Dcp1:Dcp2 autoinhibition and alleviation by Edc3

See main text for description

Table 1: Constructs used in chapter 1

Table of constructs used in main text and supplemental figures

Protein construct (Reference name)	AA Boundaries (full-length unless specified)	Solubility/ Purification Tag	Vector backbone	Organism
Dcp1:Dcp2 (catalytic core)	1-243 (Dcp2)	N-His$_6$-GB1-TEV on Dcp1	pYACYC_Duet	S. *pombe*
Dcp1:Dcp2	1-243 (Dcp2) [Y220G]	N-His$_6$-GB1-TEV on Dcp1	pYACYC_Duet	S. *pombe*
Dcp1:Dcp2	1-266 (Dcp2)	(1) N-His$_6$-GB1-TEV on Dcp1	pYACYC Duet	S. *pombe*
Dcp1:Dcp2	1-504 (Dcp2)	(1) N-His$_6$-GB1-TEV on Dcp1 (2) C-StrepII on Dcp2	pRSF	S. *pombe*
Dcp1:Dcp2	1-504 (Dcp2) [Y220G]	(1) N-His$_6$-GB1-TEV on Dcp1 (2) C-StrepII on Dcp2	pRSF	S. *pombe*
Dcp1:Dcp2 (ΔIM1)	1-504 (Dcp2) Δ267-350	(1) N-His$_6$-GB1-TEV on Dcp1 (2) C-StrepII on Dcp2	pRSF	S. *pombe*
Dcp1:Dcp2 (ΔIM2)	1-504 (Dcp2) Δ399-434	(1) N-His$_6$-GB1-TEV on Dcp1 (2) C-StrepII on Dcp2	pRSF	S. *pombe*
Dcp1:Dcp2 (ΔIM1/2)	1-504 (Dcp2) Δ267-350 Δ399-434	(1) N-His$_6$-GB1-TEV on Dcp1 (2) C-StrepII on Dcp2	pRSF	S. *pombe*
Edc3		N-His$_6$-TRX-TEV	pET30b	S. *pombe*
Edc3 (LSm)	1-94	N-His$_6$-TRX-TEV	pET30b	S. *pombe*
Edc3 (YjeF N)	188-454	N-His$_6$-TEV	pET30b	S. *pombe*

Table 2: Observed kinetic rates of constructs in Chapter 1

Table of mean k_{obs} at 1.5μM [Dcp1:Dcp2]

Complex	Mean K_{obs} (min^{-1})	Population standard deviation, σ
Dcp1:Dcp2 (1-243)	0.0892	0.00996
Dcp1:Dcp2 (1-243) Y220G	0.1361	0.0166
Dcp1:Dcp2 (1-266)	0.0658	0.0005
Dcp1:Dcp2 (1-504)	0.0114	0.0007
Dcp1:Dcp2 (1-504) Y220G	0.0722	0.0161
Dcp1:Dcp2 ΔIM1	0.0508	0.00255
Dcp1:Dcp2 ΔIM2	0.0588	0.0007
Dcp1:Dcp2 ΔIM1/IM2	0.0812	0.00505

Edc3 related k_{obs}

Complex	Mean K_{obs} (min^{-1})	Population standard deviation, σ
Dcp1:Dcp2 (1-266) + Edc3	0.2751	0.0435
Dcp1:Dcp2 (1-266) + LSm	0.1141	0.0184
Dcp1:Dcp2 (1-266) + YjeF N	0.0578	0.0085
Dcp1:Dcp2 (1-504) + Edc3	0.4116	0.0476
Dcp1:Dcp2 (1-504) + LSm	0.1077	0.00765
Dcp1:Dcp2 (1-504) + YjeF N	0.0166	0.00055
Dcp1:Dcp2 (1-504) Y220G + Edc3	0.2351	0.0035

Fig S1: Purity of proteins used in kinetic assays

Coomassie stained SDS-PAGE of purified proteins used in this study.
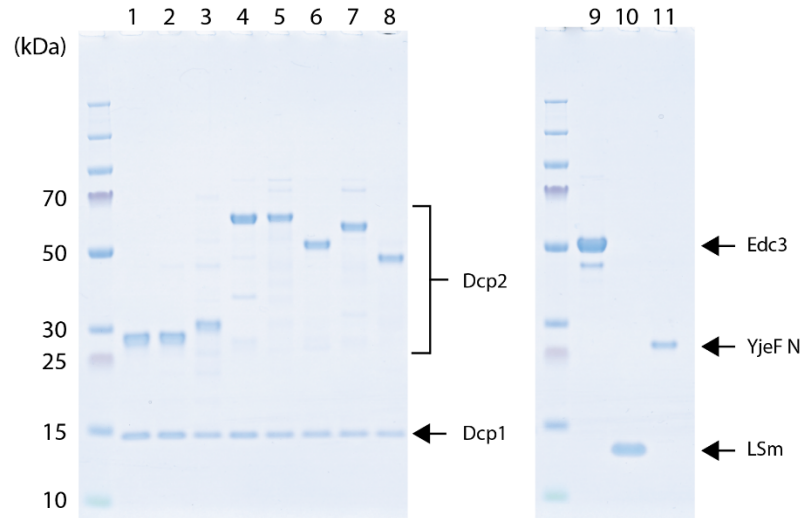
1. Dcp1:Dcp2 (1-243)
2. Dcp1:Dcp2 (1-243) Y220G
3. Dcp1:Dcp2 (1-266)
4. Dcp1:Dcp2 (1-504)
5. Dcp1:Dcp2 (1-504) Y220G
6. Dcp1:Dcp2 (1-504) ΔIM1
7. Dcp1:Dcp2 (1-504) ΔIM2
8. Dcp1:Dcp2 (1-504) ΔIM1/2
9. Edc3
10. LSm (1-94)
11. YjeF N (188-454)

Fig S2: Full alignment of fission yeast dcp2 genes


```
           1        10        20        30        40        50        60
Pombe      MSFTNATFSQVLDDLSARFILNIPAEEQSSVERLCFQIEQAHWIYEDFIRAWNDQLPSLG
Cryophilus MSFKNATFSQVLDDLSARFILNLPAEEQSSVERLCFQIEQAHWIYEDFIRAWNDQLPSLG
Octosporus MSFKNATFSQVLDDLSARFILNLPAEEQSSVERLCFQIEQAHWIYEDFIRAWNDQLPSLG
japonicus  MSFSNATFEQVLDDLSVRFLLNIPSEEKNSIERLCFQIEQAHWIYEDFIRAENDQLPSLG

                    70        80        90        100       110       120
Pombe      LRVFSARLFAHCPLLWKWSKVHEEAFDDFLRYKTRIPVRGAIMLDMSMQQCVLVKGWKAS
Cryophilus LRVFSARLFAHCPLLWKWSKVHEEAFDDFLRYKTRIPVRGAILLDKSLQRCLLVKGWKAS
Octosporus LRVFSARLFAHCPLLWKWSKVHEEAFDDFLRYKTRIPVRGAIMLDKSLQRCLLVKGWKAS
japonicus  LRVFSAKLFSHCPLLWKWSKVHEEAFDDFLRYKTRIPVRGAIMLNKEMDKCILVKGWKAS

                    130       140       150       160       170       180
Pombe      SGWGFPKGKIEKDESDVDCAHREVYEETGFDCSSPINPNEFIDMTIRGQNVRLYIIPGIS
Cryophilus SGWGFPKGKIEKDESDVDCAHREVYEETGFDLMGHINPNEYIEMTIRGQNVRLYIIPEMS
Octosporus SGWGFPKGKIEKDESDVDCAVREVYEETGFDLTGHINSNEYIDMTIRGQNVRLYIIPEMS
japonicus  SGWGFPKGKINKDEADVDCAHREVYEETGFDSTNWINEKDFIELTIRBQNHRLYIIPDLP

                    190       200       210       220       230       240
Pombe      LDTRFESRTRKEISKIEWHNLMDLPTRKKNKPQTVKNKFYMVIPFLAPLKKWIKKRNIAN
Cryophilus LDTVFESRTRKEISKIEWHMLADLPTVKKSRTDTVKNKFYMVIPFLTPLKKWIKKRNMAT
Octosporus LDTAFESRTRKEISKIEWHMLTDLPTVKKSRTDTVKNKFYMVIPFLAPLKKWIKKRNVAT
japonicus  METVFESQTRKEISKIEWHLLADLPTRKQSKSETIKNKFYMVIPFLAPLKKWIRKRLVSL

                    250       260                             270
Pombe      NTTKEKNISVDVDALASSQLLSLL....................KSST......APSDL
Cryophilus KTTKEQGISVDVDALASSQLFSLL....................KSST......AFTTI
Octosporus KTTKEQGISVDVDALASSQLLTLL....................KSST......APNVN
japonicus  EGKVQATK.......ASTQLLSLLKEPSSTATVAESANEKLLNMLRTPVQTQSQLAAPAR

                    280                 290       300       310
Pombe      ATPQPSTFPQ......PPVESH.......SSFDIKQKILHILNEGNEPKS...PIQLPP
Cryophilus PLPISQPMEQ......QAVPQI.......PHVDIKQKILTILNGNLEPKQ...SLPILP
Octosporus SLPISQPVEQ......QAVPQL.......PIVDIKQKILTLLNGNFEPKQ...PLPILP
japonicus  ALPTAISEEKLLALIRGPAAPTASTIQAPPPDDYKACILNMLNSGLSAHQSAAPMSMPP

                    320       330       340
Pombe      VSNLPLNPP.IQSSNSRLS...HDN.NSFDPFAYLG....................
Cryophilus VSSLPQQPP.RQLDHSWPN...GPSLYGFDPFAYLG....................
Octosporus VSSLPQQPP.RQVDNPWLN...GPSLYGFDPFAYLG....................
japonicus  LPVPNMHPSTTQFQSPVMNFAPTAPFFGMSPYQNFAQLHPAQGMYPVMSGPSTPQASRQM

                    350       360       370       380
Pombe      .....LDPKN.........PSASFP....RVVSQN....NMLTNKPVLNNHFQQSMYSNL
Cryophilus .....LDPHN.........PSESFP....RVAATGPPPRVASNLQFPTSFSHFPPPPPHF
Octosporus .....LDPQN.........PSESFP....RVAASGPPPRVASNLQFPDSVNYFSPPQHY
japonicus  ESTIPLDPHIQQIMSNSTDPSGTLPFTDPAIIKTNLQFANF..ELPPMRSFAAFPQAMPL

                    390       400       410       420       430       440
Pombe      LKDQN....SVQHLFAASDMPSPMELPSPSTVYHQVFYPPTSTSVSSYGLGKTPQPAYGS
Cryophilus PDYSSPYVQNRENSSPEATMPSPMDLPSPKTVYHVFHPPTTRSVQSFKPEQKR.EVYND
Octosporus LEHRNPYVQNRENFSPEAAMPSPMDLPSPKTVYHQVFHPPSTRSVQSFQPEQMR.EVYND
japonicus  PSQSD....I.......SSVPTPGFLPSPSTMYHQVFHPPSMTAVTTFEKLEKN.A....

                    450       460       470       480       490
Pombe      SSPYVNG....HQIQQISSLPPPQSQTQFLARNSDNSGQSYNSEGDSNSKRLLSMLSQQD
Cryophilus THPSLNMQIPAYTNQMIDPFPTIQHRQK.MKSFHDESSLN.RSQNDASSRKLLDMISSPK
Octosporus THPKLNIQNPPYTNHTVDPFAVIQQQQQPINNVHDEPTVN.RLQNNASSRKLLDMISSPK
japonicus  ...........................................ASKGNSDNNTQERIVGHKK
```
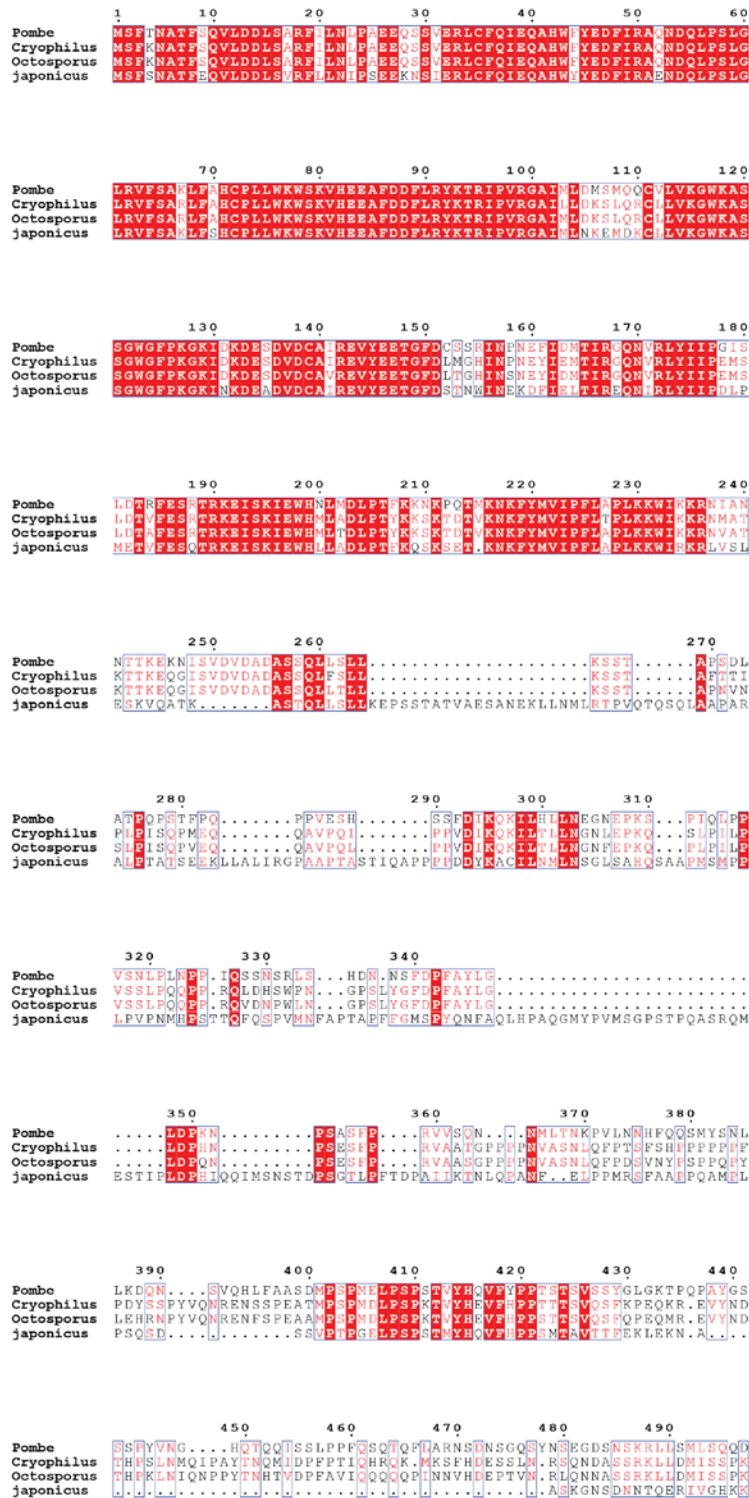
31

```
                500        510        520        530        540        550
Pombe       TTPSSSTLSKEANVQLANLFLTPNSLETKKFSDNSQGEEISDNLHGESCNNPNAN.SVHS
Cryophilus  PAHSFPTSRDGSSQKLVNVLLS.....TPKPDNNVNVPEKAETINKPAVSKLSTQPSSHS
Octosporus  PARSLPASRDGSSQKLVNILLG.....TDNSDVNGSIPNKTETVHKPIVSNPPAQPSSHS
japonicus   QKPFTDDSRE...........A...........EQKSKGQASKQDSNPSV..

                560        570        580        590        600        610
Pombe       AQLLQALLHPSATETKEETPKKTSDSLSLLTLLKSGLPTPANDLQNKSQNNERKASSQVK
Cryophilus  EQLLEALLNPSTF....SLPSNLSSSAITNNELNGAQPLETKAS........PKPKSRGV
Octosporus  EHLLQALLNPTSF....AKPS....SAAEKNEP.GNQPLEAKPL........PKPKSTIP
japonicus   .QLLNILTKAAEL....S......LNEKRQVKSKSVAEASRTQ........STKLQ

                620        630        640        650        660        670
Pombe       ELEVKNYSKSTDLLKKTLRIPRNDEPLEAANQFDLLKVSPQQKSEVPPKRNELSQSKLKN
Cryophilus  QLDAKSEAEDNDATRK.....KKTEALKNV....LI......GSSVQPMANTSTEVPIQL
Octosporus  QSDASDKPKSVEFTRV.....ENTEALKNV....LI......GRSIQPTTADSKEVMTEH
japonicus   PTVSSPYKTSNQFSKGAT...KKKSP.........RVSTPKRRNGLSGGKQTF

                680        690        700        710        720        730
Pombe       RKKKENSETNKNHVDMSPGFVKILKRSPLALQKKEDTQESDFKGSDDHFLSYLQSVVSSN
Cryophilus  TAKADLSIQEPTTPGTFPGSVKILKRAPSSEVDDKSKQTV.EKSSDEYFLSYLQSVVPWI
Octosporus  GIKTDGSTQKPATPGFFPGSVKILKRAPSAELEGKNHPTV.GKSSDEYFLSYLQDVVRNQ
japonicus   SPKPVSPFGNK...ELTPTNVKILKREPVSKTRT........TANDSLFMDYLKSVVTNT

                740
Pombe       S.NGLH
Cryophilus  IQLIVK
Octosporus  KR...T
japonicus   NPTGAP
```

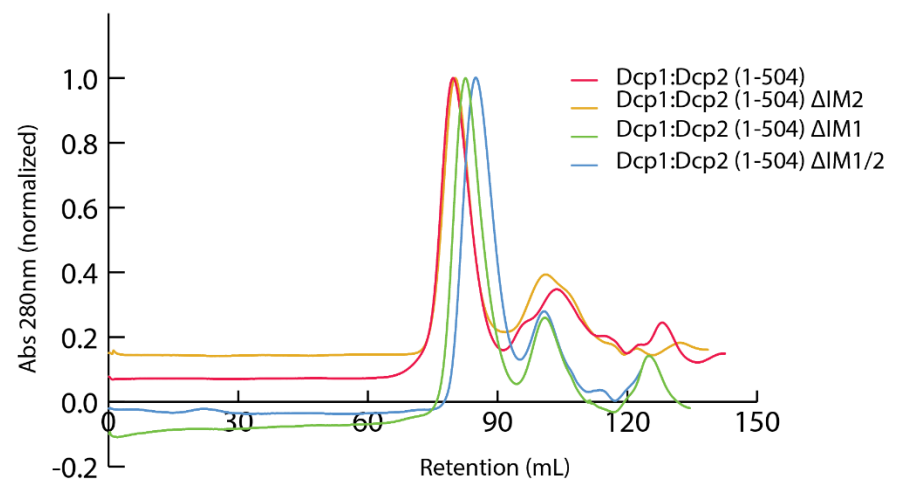Fig S2B: Size-exclusion chromotagraphy of dcp1:dcp2(1-504) internal deletions

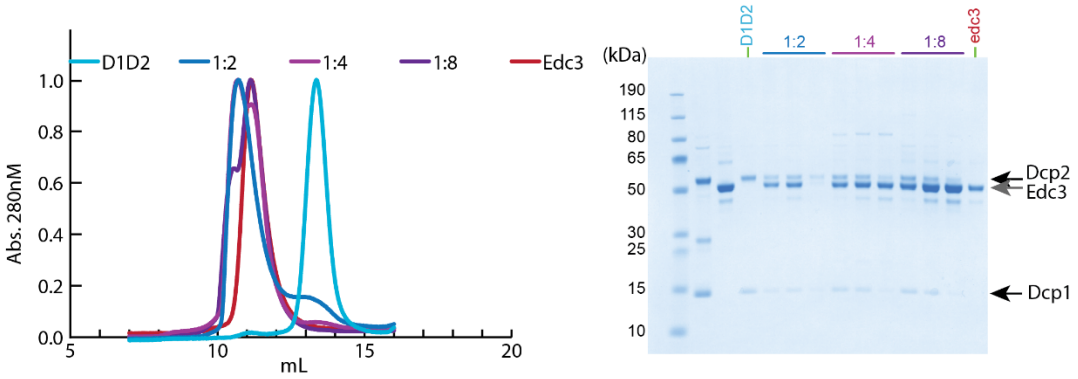Fig S3: edc3 complex formation with Dcp1:Dcp2(1-504)

Fig S4: Cap-occluded and Y220 related

S. pombe   FK.....KNKPQTMKNKFYMVIPFLAPLKKWIKKRNIANNTTKEKNISVDVDADASSQLL
H. sapien   HRNDMTPKSKLGLAPNKFFMAIPFIRPLRDWLSRRFGDSSDSDNGFSSTGSTPAKPTVE.
M. musculus   HRNDMTPKSKLGLAPNKFFMAIPFIRPLRDWLSRRFGDSSDSDNGFSSAGSTPARPTVE.
X. laevis   HKNDMTPKSKLGMAPNKFFMVIPFIRDLRDWLHRKFGDSSDSDNGFASAGSTPSKPNAE.
D. rerio   HRNDMTPKSKLGLAPNKFFMAIPFMRQLKEWIAKHKGESTSSDEDFASNGSTPCKPFRS.
D. melanogaster NKNDAISKAKLGKTSNSFFMIMPFVKRLKKWVNDRKAGIEPRRRKFSGQQSPKAASPTNM
S. cerevisiae ..|.....TMYKS..NIKYYLINSMMRPLSMWLRHQRQIKNEDQLKSYAEEQLKLLLG...
K. lactis   ..|.....AITKNEGSAKVFLVNSMIRPLSLYVKNEKRAKDENKLKLYAEEHLKSILG...
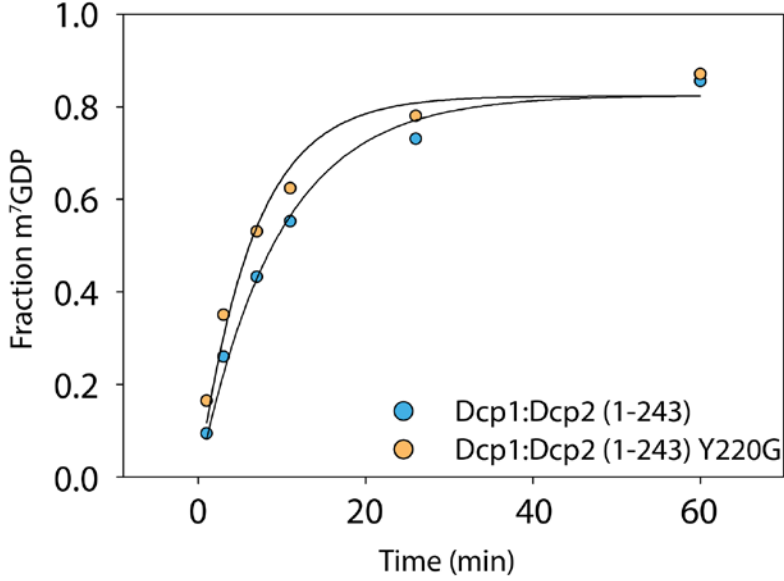
Ω

Fraction m^{7}GDP vs Time (min)
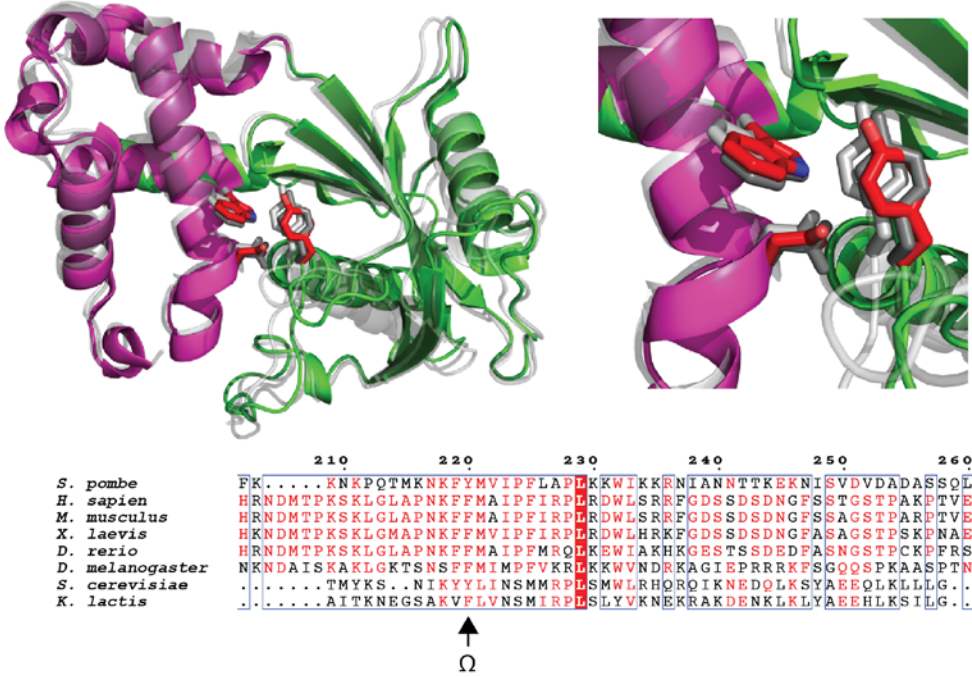
● Dcp1:Dcp2 (1-243)
● Dcp1:Dcp2 (1-243) Y220G
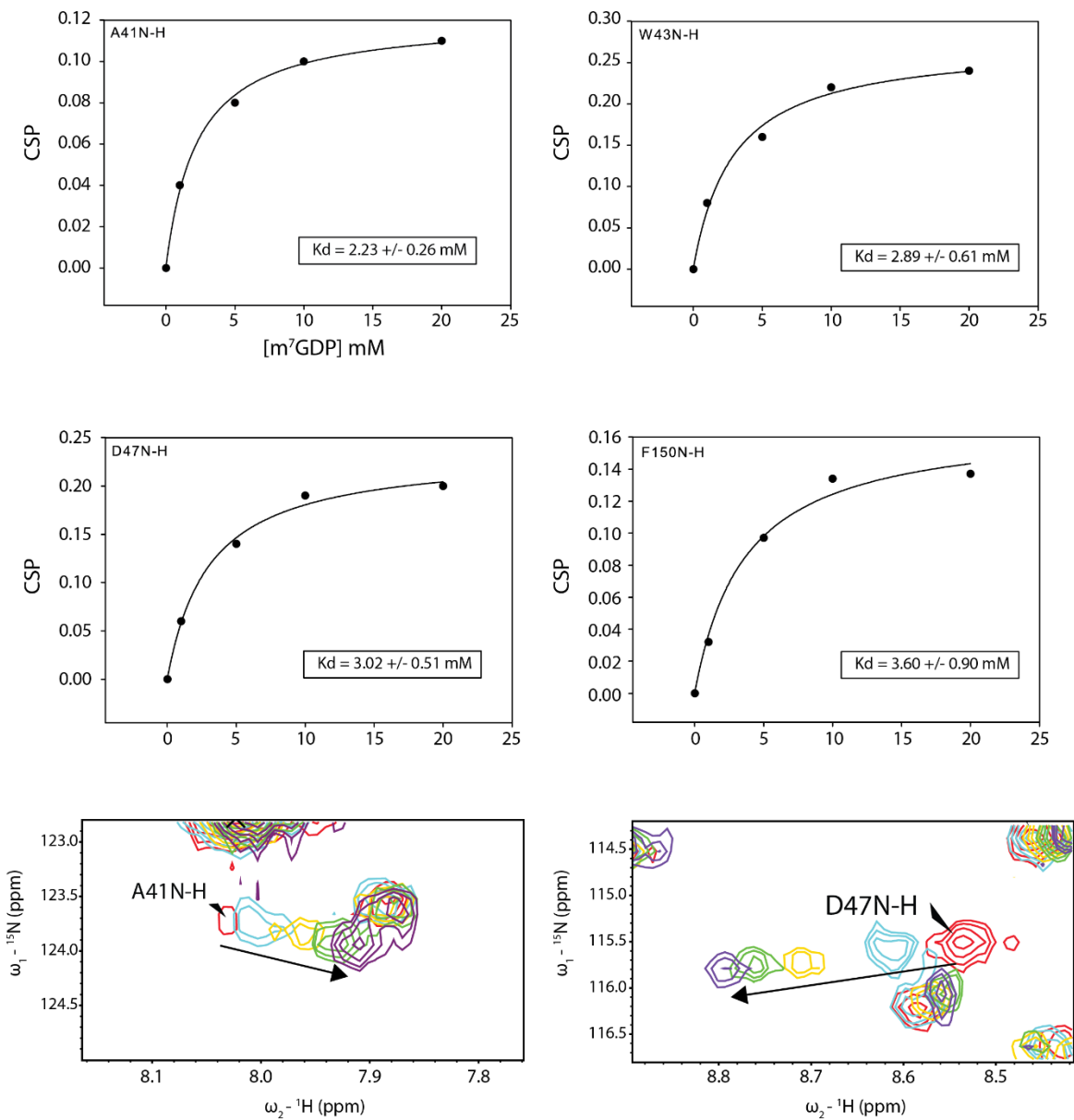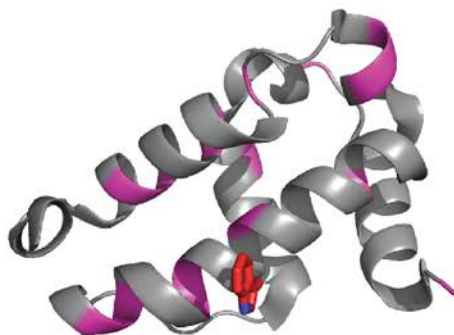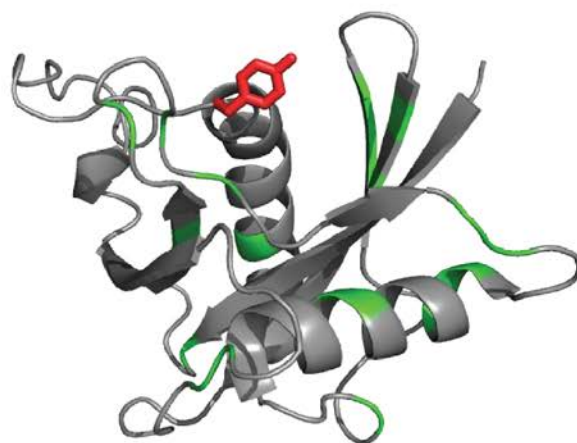
Fig S5: Cap-occluded and Y220 related

Fig S6: Amide resonances that reappear in both W43A and Y220G Dcp2 mutants

Regulatory domain:

Catalytic domain:

SUPPLEMENTARY FIGURE LEGENDS

Table 1: List of plasmids used in this study.

Table2: List of observed decapping rates for the various constructs from table 1 and in complex with Edc3.

Fig S1: SDS-PAGE showing purity of the constructs from table 1 as used in biochemical studies.

Fig S2: (A) Full sequence alignment of fission yeast Dcp2 proteins. (B) Size-exclusion chromatograms of the various Dcp1:Dcp2 constructs used to map the regions contributing to the autoinhibition.

Fig S3: Analytical S200 size-exclusion chromatograms of Dcp1:Dcp2 (1-504) or Edc3 alone or in complex with increasing excess of Edc3. SDS-PAGE of fractions from the center of the peaks from a given sizing chromatogram.

Fig S4: (A) Alignment of ATP bound form of Dcp2 (2QKM) colored as in Fig4 A with 5QK1 and 5J3Y showing the common cap-occluded conformation, where W43, D47 and Y220 make contacts. Alignment below show that Y220 is a conserved aromatic. (B) Plot of the fraction m^{7}GDP (product) versus time for either WT or Y220G Dcp1:Dcp2 (1-243) showing slightly increased activity.

Fig S5: Plots of CSP (ppm) versus m^{7}GDP concentration from HSQC cap binding titration experiments. Residues used to determine cap binding affinity with corresponding fits and representative cross-peaks show shifts with increasing [m^{7}GDP].

Fig S6: Visualization of the common backbone amide resonances that reappear in both W43A and Y220G Dcp2 mutants. Top shows RD colored in magenta and bottom shows CD colored in green. W43 and Y220 side-chains are shown as stick rendition colored in red. Peaks in both return from domains distal from the mutation.

APPENDIX

IM2 binding characterization by fluorescence anisotropy

Since the peptide derived from the Dcp2 IM2 (399-434) is proline rich in nature and shares some similarities with the known sequence that binds the aromatic quartet of Dcp1, we thought that the IM2 might bind in a similar manner. We decided to make use of a fluorescence anisotropy based assay to measure the binding of a of an IAEDANS-IM2 peptide to the structured Dcp1:Dcp2 (1-243) catalytic core. Using the same protocol as detailed in (Borja et al. 2011), we first measured the K_d of IM2 to Dcp1:Dcp2 and noticed that the binding was weaker than the canonical Edc1-like peptides, i.e. PNRC2 (**Fig 1A**). We then tested if the binding mode was similar to Edc1 by making known mutations to the Dcp1 binding surface that are known to increase the K_d. While the mutations impacted Edc1/PNRC2 as expected (Lai et al. 2012), IM2 did not show the same behavior suggesting that it engages Dcp1 in a different manner (**Fig 1A**). For Edc1-like peptides, it is known that Dcp1 provides the majority of the free-energy of binding, such that removal of the catalytic domain (CD) of Dcp2 (95-243) should have no effect on the binding. For PNRC2, a canonical co-activator peptide, we see that the absence of the CD does not affect the binding; however, for IM2 we appear to lose a majority of the binding energy **(Fig 1B)**. Following up on this behavior, we asked if the CD alone is sufficient for IM2 binding. In fact, IM2 binds to the CD alone with a lower apparent K_d (**Fig 1C**). The IM2 peptide also binds to Dcp2 (1-243) in the absence of Dcp1, but with a higher affinity than Dcp1:Dcp2 (1-243), implicating a role of Dcp1 in the binding either directly or through biasing of the Dcp2 CD binding interface **(Fig 1D,E)**. Finally, a D47A mutation that is known to quench the Dcp2 milisecond/microsecond solution dynamics lowers the K_d for the IM2 peptide. We interpret this change in binding to an increase in the accessibility of a CD binding surface. Consistent with the CD being more accessible, we also see an increase in the anisotropy of the fully bound peptide state suggesting that the enzyme is in a different less compact conformation than the WT protein **(Fig 1D)**. Further work needs

to be carried out to further map the interactions of the IM2 both biophysically and structurally to better understand the mechanism of decapping autoinhibition.

Characterization of *S. cerevisae* specific IDR conservation

In work from the Jacobson lab (He and Jacobson 2015), they had identified *in vivo* that there was a region of *S. cerevisiae* Dcp2 that appeared to negatively regulate the activity within certain strain backgrounds. Specifically, mRNAs that are modulated by Edc3 were stabilized in an Edc3 null strain, but this stabilization was bypassed when a truncated Dcp2 was introduced in place of the endogenous enzyme. We decided to see if the negative regulation by this identified region was a direct effect on the decapping activity as interpreted in the published findings. We designed and purified a Dcp1:Dcp2 (1-451) coexpression construct and were able to after sizing on a Superdex 200 get a reasonably biochemically pure sample (**Fig 2A,B**). We compared the activity of this C-terminally extended complex with the catalytic core, Dcp1:Dcp2 (1-245). We measured only a modest effect on the decapping activity of the complex in the context of the longer construct on our 33mer RNA substrate (**Fig 2C**). We cannot rule out that on a longer or more specific RNA the effect could differ. The fold-change in the rate at two different concentrations was less than 2-fold. Based on further information from *pombe* where a region of the IDR binds back onto Dcp1 (Wurm et al. 2016), it would be informative to see if there is potentially an inhibition to Edc1 dependent activation. We cannot rule out if it could compete for Dcp1 binding or if we are missing additional factors that are present in the cellular environment. Further follow-up experiments would need to be performed to fully rule out if an additional cofactor is required.

Fig. 1: S. pombe Dcp2 (399-434) binds to Dcp1 and the CD of Dcp2
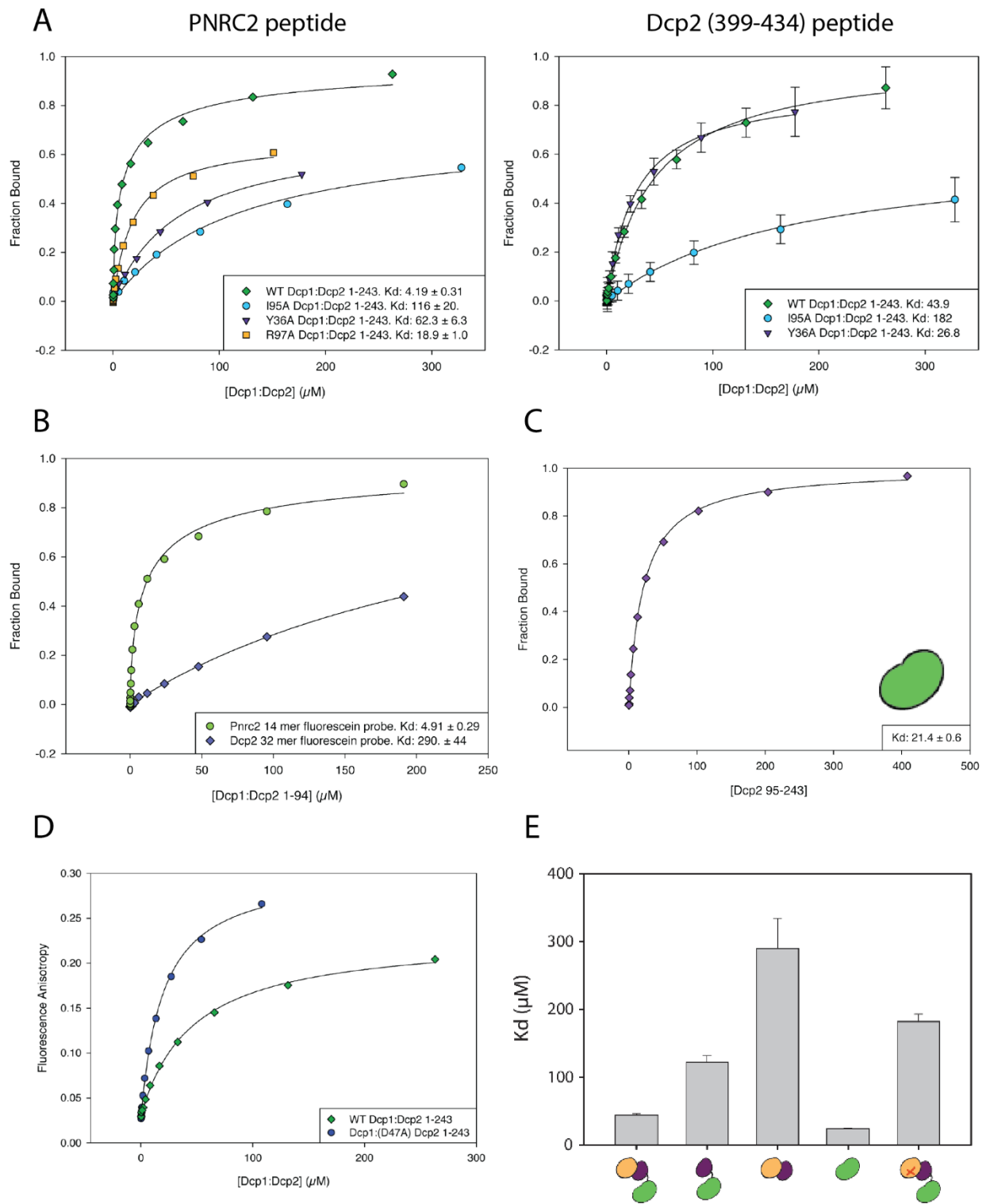
Fig. 2: C-terminally extended S. cerevisiae Dcp1:Dcp2 does not appear to be autoinhibited

	k_{obs} (min^{-1})	F_{ep}
d1d2(1-245) 1uM	0.71	0.84
d1d2(1-451) 1uM	0.50	0.73
d1d2(1-245) 13uM	2.54	0.87
d1d2(1-451) 13uM	1.63	0.85
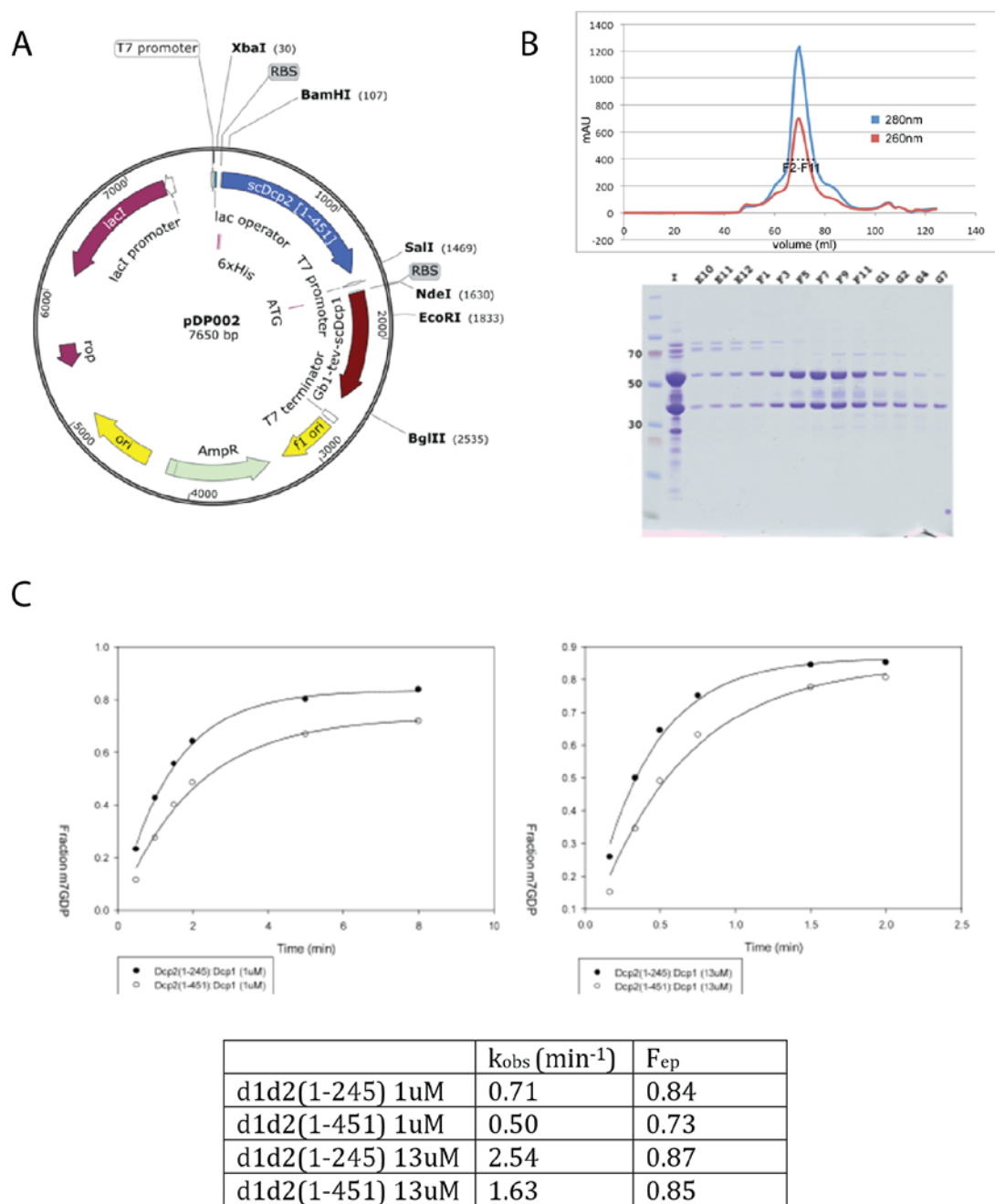
APPENDIX FIGURE LEGENDS

Fig 1: (A) Fluorescence polarization derived binding curves for either labeled PNRC2 (15mer) or IM2

probes and titrating WT Dcp1:Dcp2 (1-243) and Dcp1 point mutant complexes. (B) Binding curves of

PNRC2 and IM2 with Dcp1:Dcp2 (1-94), which lacks the CD. (C) Binding curve for IM2 with the CD. (D)

Fluorescence anisotropy of IM2 bound to either WT or Dcp1:Dcp2 (1-243) D47A. (E) Bar graph of the

measured K_ds for various combinations and mutants of Dcp1:Dcp2 for IM2

Fig 2: Plasmid map of vector used to overexpress *S. cerevisiae* GB1-Dcp1:Dcp2 (1-451). (B) Size-

exclusion chromatogram and SDS-PAGE of the major peak containing the complex. (C) Fraction m^7GDP

versus time for the catalytic core (1-245) and C-terminal extension (1-451) at 1 µM and 13 µM. Table

shows the empirically determined rates.

Chapter 2

An optimized activated decapping complex from *S. pombe* as an ideal TAP replacement for TSS mapping

David R. Paquette[1], Jeffrey S. Mugridge[2], David E. Weinberg and John D. Gross[1,2]

[1] Integrative Program in Quantitative Biology, Biophysics division, University of California, San Francisco, California, 94158

[2] Department of Pharmaceutical Chemistry, University of California, San Francisco, California, 94158

[3]Department of Cellular and Molecular Pharmacology, and in the [4]Sandler Faculty Fellows Program, University of California, San Francisco, California 94158

ABSTRACT

Changes in the 5' leader can have profound effects on a gene's translational efficiency with little affect to its transcript abundance. Additionally, this feature can result in disease or cell-specific changes. Several sequencing based methods to accurately map this 5' leader heterogeneity by identifying the first-transcribed nucleotide rely on the enzymatic removal of the eukaryotic cap structure. We have designed and validated a superior tobacco acid pyrophosphatase (TAP) replacement that can be purified from *E. coli* overexpression using standard biochemical methods. This super decapping enzyme is three-orders of magnitude more catalytically efficient at m^7G cap removal, it is effective under multiple-turnover conditions, and the downstream product, monophosphorylated RNA, is suitable for standard RNA coning methods. This enzyme is a better replacement than the current marketed use of RppH.

INTRODUCTION

It has been appreciated for at least the past couple of decades that features in the 5'UTR can impact mRNA stability and translational efficiency (Kapranov 2009). It was not until even more recently that inexpensive and rapid deep-sequencing from the likes of likes of Illumina that we have come to appreciate the complexity and heterogeneity in the 5'UTRs of many genes (Pelechano et al. 2013; Gu et al. 2012; Farcas et al. 2012). From simple single-cell eukaryotes to complex multicellular humans, mapping of the 5'UTR space has rewritten our understanding of genetic plasticity and its impact on human disease. Researchers studying *S. cer*evisiae have determined that variation in the 5'UTR of genes can result in translational efficiencies that can vary by as little as 3-fold to greater than 100-fold with little alteration to mRNA levels (Rojas-Duran and Gilbert 2012). Importantly, this effect of alternative 5'UTRs has been shown to be conserved in humans (Calvo et al. 2009). In multicellular organisms that contain specialized cell-types, this heterogeneity can vary amongst cells and can also be altered between non-turmoural and tumoural cells resulting in quantitative and qualitative differences in proteins between cell-types (Dieudonné et al. 2015)

In the human disease frontier, alterations in the 5'UTRs have been shown to contribute to a variety of disease states (review). One of the earliest descriptions of a change in the 5'UTR of a gene resulted in the overexpression of the oncoprotein MDM2 in a number of human tumors (Brown et al. 1999; Jin et al. 2003). The 5' leader sequences did not alter stability or amount of mRNA but rather the translational efficiency by the excision of repressive upstream ORFs. In addition to differential splicing and alternative TSSs, mutations that introduce or disrupt a uORF have been found to cause at least three human diseases: Thrombocythemia, Melanoma and Marie Unna hereditary hypotrichosis; and several more have been predicted (Calvo et al. 2009)

With the new lens of the large impact that 5' leader sequences can have on translational efficiency, it is important to be able to reproducibly map the TSS. A variety of RNA-seq methods have been innovated to accurately map the first transcribed nucleotide (source listing methods). Many of these methods rely on the enzymatic removal of the m^{7}G cap structure and subsequent ligation of an adapter sequence to the reactive 5' monophosphate. Historically, this reaction was performed by using Tobacco Acid Pyrophosphatase (TAP); however, TAP production was discontinued. RppH has been promoted as an alternative reagent, but its specific activity is towards triphosphorylated, uncapped bacterial RNAs and not eukaryotic 5' capped RNAs (Deana et al. 2008). With the increased importance of accurately determining TSSs in WT and diseased states, it would be ideal to have a more specific reagent for mapping the first transcribed nucleotide.

Here, we describe the use of the *S. pombe* activated decapping complex as a TAP replacement for mapping TSSs. The complex is recombinantly expressed using standard *E. coli* protein production systems from a single, iPTG inducible plasmid. Additionally, the complex is readily purified using a Ni-NTA affinity column, followed by removal of a solubility tag and purification to homogeneity by size-exclusion chromatography. The purified complex was compared side by side with RppH and showed three-orders of magnitude greater catalytic efficiency in an *in vitro* decapping assay. Furthermore, the enzyme showed greater activity in the more sequencing relevant scenario of total yeast RNA. We also demonstrate that the 5' end of the enzymatically treated total RNA can be cloned onto using standard methods. This easy to produce and highly active decapping complex should make for a superior TAP replacement for both small scale and large scale sequencing methods for mapping the 5' end of RNAs more readily.

RESULTS

Design and construction of the activated decapping complex (super decapper)

The fungal decapping enzyme forms a stable obligate heterodimer that can be recombinantly expressed with high yields. Furthermore, its catalytic enhancement by the conserved YAG motif containing Edc1-like activators is well described both biochemically and structurally (Borja et al. 2011; Lai et al. 2012; Wurm et al. 2016; Schütz et al. 2017). We set out to design a constitutively activated super decapping complex consisting of the core Dcp1:Dcp2 stably bound by the minimal sequence of the Edc1 coactivator. We designed an IPTG inducible single plasmid expression system consisting of Gb1-tev-Dcp1 and a hexahistadine-TRX-tev-Edc1-GSlinker-Dcp2, where the conserved activation sequence of Edc1, residues 155-186, are stably fused to the N-terminus of Dcp2 (**Fig 1A**). We designed the Edc1-Dcp2 fusion construct with two linker lengths, 5aa and 10aa, since we were unsure which would better facilitate the endogenous contacts. Both the 5aa and 10aa linker complexes were eluted off Ni-NTA as relatively pure complexes and both sets of solubility tags were fully cleaved off (**Fig 1B**). Size-exclusion chromatography revealed that the 5aa linker complex ran as a heterotetramer, whereas the 10aa linker complex ran as a mixture of the heterotetramer/heterodimer where the heterodimer was predominant **(Fig 1C).** For sample homogeneity and reproducibility, we decided to use the 10aa linker construct, specifically the lower molecular weight heterodimer peak, for all further experiments.

Super decapper enzyme is more catalytically efficient than RppH

RppH is currently the recommended, most readily available enzyme for 5' m^7G cap removal from eukaryotic RNA. We set out to benchmark our constitutively activated super decapping enzyme against the industry standard. We performed a series of decapping assays at varying concentrations of enzyme

to determine K_m and V_{max} under our assay conditions for both enzymes. Unfortunately, we were unable

to reach the V_{max} for RppH, so we instead compared the catalytic efficiencies of the enzymes (i.e. the slope

of the linear fit, which is equivalent to V_{max}/K_m). The super decapping enzyme is clearly the better reagent

for removing the cap structure from eukaryotic RNA, as judged by catalytic efficiency. RppH in buffer

conditions recommended by NEB is nowhere near as efficient (**Fig 2A/B).** The above results were under

single-turnover conditions where enzyme is in excess of substrate and at 4 °C to ensure we could more

accurately measure the rates of the supper decapping enzyme. We also wanted to compare the

effectiveness of the enzymes under more relevant conditions were substrate is in excess of enzyme (**Fig

2C**) In order to do this, we performed the same assay in the presence of excess *pombe* total RNA. We

initiated the reactions side by side at 37 °C, as recommended by NEB since we used readily available RppH

for this comparison. The super decapping enzyme was almost too fast to measure by manual pipetting

even at relatively low enzyme concentrations (**Fig S1)**. Despite the difficulties in accurately measuring the

observed rates, it is apparent that the super decapping enzyme is a better enzyme for specifically

removing the eukaryotic cap as compared to RppH.

Downstream RNA cloning applications demonstration

While the super decapping enzyme is a more catalytically efficient enzyme for eukaryotic RNA

5'cap removal, we wanted to ensure that the treated RNA was suitable to downstream cloning

applications. We decided to use a previously reported upon assay (Blewett et al. 2011; Wang et al. 2013)

as verification that: (1) we could detect the decapping of an mRNA from bulk *pombe* purified RNA, and (2)

we could use a cloning method that relies on a 5' monophosphate and an intact first-transcribed

nucleotide. We treated three 20 µg samples of purified *pombe* total RNA (**Fig S2A)** in the presence of 1

µM super decapping complex and stopped the reaction by addition of phenol-chloroform after 15, 30 and

60 minutes. The quenched reactions were ethanol precipitated and taken through the splinted-ligation procedure as detailed in (**Fig 3A**). The final PCR products of the splinted-ligation reaction were visualized by sybr gold in a 2% TAE agarose gel (**Fig 3B**). All three samples had a similar amount of total cDNA product, but the splinted-ligation cDNA product showed a time-dependent increase; consistent with an increase in the amount of decapped and thus ligateable RNA transcript. While we only show one cloning method here, any protocol that relies on a 5' monophosphate should be well suited.

DISCUSSION

With our increased understanding of the large effects that alternate 5' transcription leaders can have on protein expression it had been increasingly important to map this space in regards to humans disease. The methods to identify all the transcription start sites rely on the removal of the eukaryotic cap structure first; followed by deep-sequencing. Here we have detailed the construction of a constitutively activated *pombe* decapping complex consisting of Dcp1 and Dcp2 with the conserved co-activator sequence from Edc1 fused to its N-terminus. We have dubbed this complex the super decapping enzyme. This complex is an ideal replacement for Tobacco Acid Pyrophosphatase that can be overexpressed in standard recombinant *E. coli* expression systems and is readily purifiable using standard methods. Furthermore, we have demonstrated that the super decapping enzyme is at least three orders of magnitude more catalytically efficient than RppH. Finally, we have shown that the enzyme can decap an endogenous mRNA from *pombe* total RNA, and the resulting product is suitable for RNA cloning methods that make use of a 5' monophosphate.

MATERIALS & METHODS

Plasmid Construction and Protein Purification

The single chain his-TRX-tev-SpEdc1(155-186)-(GGGGS)2-SpDcp2(1-243) construct (Edc1-Dcp2 with 10aa linker) was obtained as an E. coli codon-optimized DNA sequence from Integrated DNA Technologies (his is hexahistadine affinity tag, TRX is thioredoxin solubility tag, tev is Tobacco Etch Virus protease cleavage site). The single chain Edc1-Dcp2 fragment was cloned into MCS1 of a Novagen pETduet expression vector containing GB1-tev- SpDcp1(1-127) in MCS2. The Edc1-Dcp2/Dcp1 construct was expressed in E. coli BL21-star DE3 cells (Invitrogen) grown in LB media with an induction time of 18 hours at 20 °C. Cells were harvested at 5000g, lysed by sonication, and clarified at 16,000g in lysis buffer (50 mM sodium phosphate pH 7.5, 300 mM sodium chloride, 10 mM imidazole, 5% glycerol, 10 mM 2-mercaptoethanol, Roche EDTA-free protease inhibitor cocktail). The protein complex was purified by Ni-NTA affinity chromatography and solubility/affinity tags were cleaved by treatment with TEV protease overnight at room temperature. Edc1-Dcp2/Dcp1 complex was separated from cleaved tags by size exclusion chromatography on a GE Superdex 75 16/60 gel filtration column in SEC buffer (50 mM HEPES pH 7, 150 mM NaCl, 1 mM DTT). The purified complex was concentrated to 14 mg/mL and flash frozen in liquid nitrogen. RppH was purified as described in (Deana et al. 2008), with some modifications. Standard BL21(DE3) NEB cells were used and additional size-exclusion chromatography step was added to separate the enzyme from a nucleic acid contaminant.

***In vitro decapping* assays**

Single-turnover in vitro decapping assays were carried out as previously described (Jones et al. 2008). A ^{32}P-labeled 355-mer (A15 tail) RNA substrate derived from MFA2 was used for all the decapping assays.

Reactions were initiated by mixture of 30 μL 3× protein solution with 60 μL 1.5× RNA solution; final dcp1:dcp2 or RppH concentration was varied and the final RNA concentration was <100 pM. For experiments with *S. pombe* total RNA, the 20 μg of total RNA was spiked in right as the reactions were being initiated. Reactions with yeast total RNA were carried out in a 37 °C heat block, whereas the other reactions were carried out at 4 °C. Store bought RppH from NEB was used for the total RNA experiments. Time points were quenched by addition to excess EDTA, TLC was used to separate the RNA from the product m^{7}GDP, and the fraction decapped was quantified with a GE Typhoon scanner and ImageQuant software. Fraction m^{7}GDP versus time were plotted and fit to a 1st-order exponential to obtain k_{obs}; when the kinetics were too slow to obtain reliable exponential fits, k_{obs} was obtained from a linear fit of the initial rates by division of the slope by the empirically derived endpoint. Catalytic efficiency (V_{max}/K_m) were obtained as detailed in (Jones et al. 2008) for the super decapping enzyme or by taking the slope from a linear fit for RppH.

Splinted-ligation RT-PCR

Splinted-ligation RT-PCR was carried out as previously described (Wang et al. 2013), but with some modifications from the Coller Lab methods online. Briefly, 20 μg of purified yeast total RNA was treated with 1uM of super decapping complex at 37 °C for 15, 30 or 60 minutes before the reaction was quenched and the RNA purified by Phenol-Chloroform extraction. 3M NaOAc, glycoblue and ~1.5-fold of 100% ice-cold ethanol were added to the RNA solution. The RNA was precipitated at -20 °C overnight. RNA was recovered by centrifugation at 14,000 rpm for 30min at 4 °C and the pellet was washed with 70% ethanol followed by brief vortexing. The RNA was pelleted at 14,000 rpm for 5min at 4 °C and the pellet was dried briefly before being resuspended in 13.3uL of nuclease-free water (Ambion). Concentration was measured on a nano-drop. 15 μg of decapped total RNA was incubated with gene specific DNA splint and

RNA anchor in DNA ligase buffer with T4 DNA ligase. The reaction was incubated at room temperature

for 16 hours. The next morning, the sample was DNaseI treated for 3 hours at 37 °C followed by Phenol-

chloroform extraction and ethanol precipitation for 2 hours to overnight.   RNA was pelleted and

resuspended as previously and the RNA concentration was determined. 5 µg of this RNA was reverse

transcribed using a gene-specific primer and the Invitrogen Superscript III first-strand synthesis kit. 2uL

of the RT reaction was used for PCR with primers either for total transcript or splinted-ligated product

amplification. Products were resolved on a 2% TAE agarose gel and stained with sybr gold.

Table 1: List of oligos used for Fig 3

List of oligos used for splinted-ligation RT-PCR for *rps23* mRNA (Fig 3):

RNA oligo	
RNA Anchor	GCUGAUGGCGAUGAAUGAACACUGCGUUUGCUGGCUUUGAUG
DNA oligo	
Anchor forward	GCTGATGGCGATGAATGAACACTGC
Rps23 Splint	AACTGTGTGTTGGATTTAGGTGAACTTAGCCATCAAAGCCAGCAAACGCAGTGTTCAT TCATCGCCATCAGC
Rps23 RT primer	CAAACAACTTTTTATTAAGTTCGT
Rps23 forward	AACCCGCAGGATTGAACG
Rps23 reverse	TGATGACTTGTAAGCGGTTCCC

Fig 1: Design and purification of the super decapping enzyme
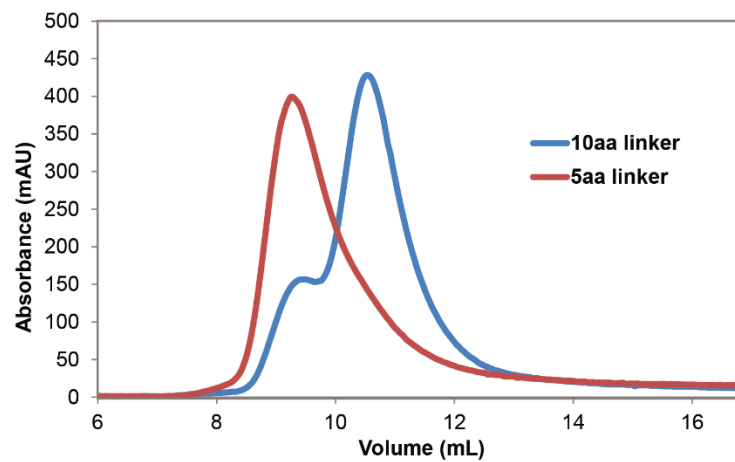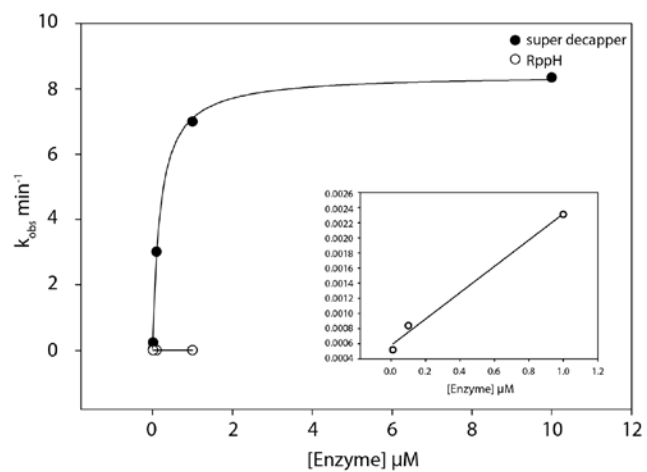
A.
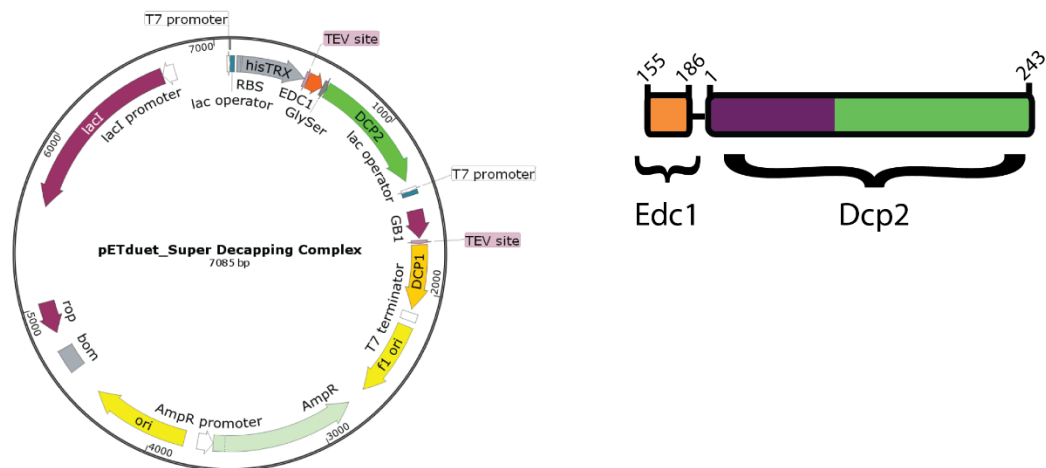
B.
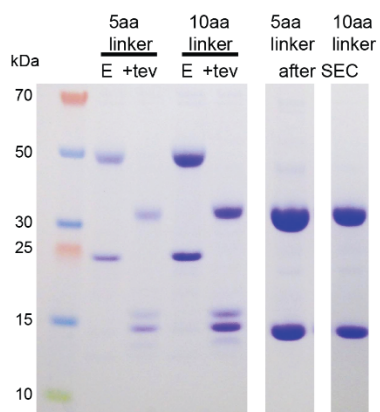
C.

Fig 2: Super decapping enzyme is more catalytically efficient than RppH

A.

B.

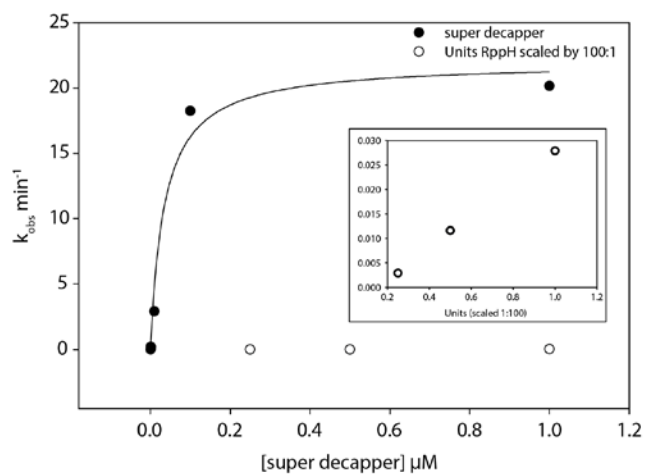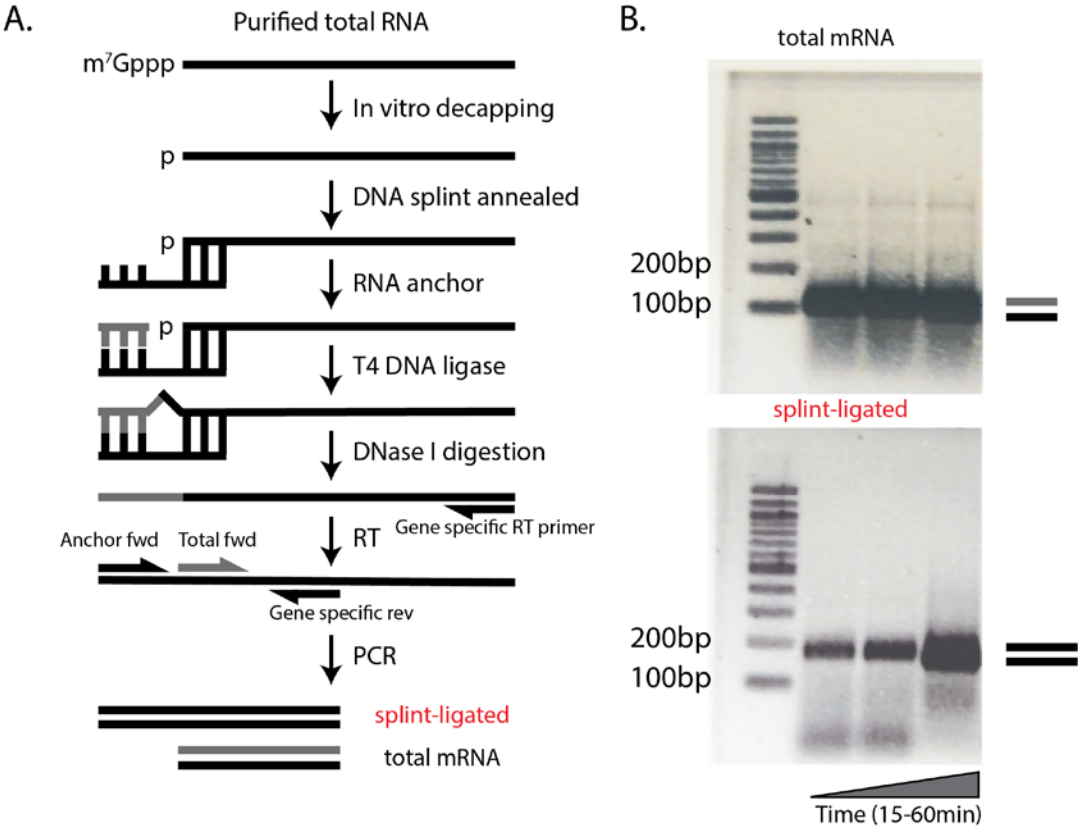| Enzyme	Catalytic Efficiency (μM^{-1} min^{-1})
RppH	0.0017
Super decapper	44.4

C.

Fig 3: Super decapping enzyme product RNA is suitable for 5' end cloning

MAIN FIGURE LEGENDS

Fig 1: (A) *E. coli* overexpression plasmid constructed from a pETduet with Gb1-tev-Dcp1 in MCS2 and the engineered single-chain his-TRX-tev-Edc1(155-186)-(GGGGS)$_X$-Dcp2(1-243) in MCS1, where X is 1 or 2. (B) Protein diagram of the super decapping dcp2 enzyme consisting of the minimal activating peptide fragment of Edc1 (orange) fused to the N-terminus of the Dcp2 catalytic core (Regulatory domain is purple and catalytic domain is green). Line connecting two proteins is the GlySer linker, either 5aa or 10aa in length. (C) Coomassie stained SDS-PAGE of both the 5aa and 10aa containing super decapping complexes after Ni-NTA elution and post-TEV cleavage. The last two strips show the final purified protein complexes post size-exclusion chromatography at the final concentration before flash freezing. (D) Analytical size exclusion chromatograms of both 5aa and 10aa complexes on a GE Superdex 75 10/300 GL column. The 10aa complex (blue trace) runs as both a higher-order hetero-tetramer as well as the expected heterodimer of Dcp1:(Edc1)-Dcp2. The 5aa complex (red trace) runs as a single hetero-tetramer of Dcp1:(Edc1)-Dcp2.

Fig 2: (A) Plot of enzyme concentration versus k_{obs} in a single-turnover decapping kinetics assay comparing the activity of the super decapping complex (filled circles) and RppH (open circles) at 4 °C. RppH is fit linearly to extract V_{max}/K_m and the super decapping complex is fit to a quadratic to obtain both V_{max} and K_m. (B Table comparing the catalytic efficiencies, defined as V_{max}/K_m, of RppH and the super decapping complex as determined from the fits in B. (C) Plot of enzyme concentration (or 100:1 Units scaled for NEB RppH) versus k_{obs} in a multiple turnover decapping kinetics assay where unlabeled purified *pombe* total RNA is in excess and the monitored p^{32} capped RNA is at the same concentration as the single-turnover

experiments. Super decapping complex (filled circles) is fit to the same equation as in B. NEB RppH (open

circles) is unfit since concentration is unknown and is reported as Units instead.

Fig 3: (A) Schematic of the splinted-ligation RT-PCR assay used to monitor the decapping of endogenous

RNA. Only decapped RNA will produce a PCR product using the anchor forward primer. The RNA anchor

oligo will only ligate if there is a free 5' monophosphate, such as the one produced upon decapping. (B)

2% TAE-agarose gels of the splinted-ligation reaction following treatment with 1 µM super decapping

complex for 15, 30 or 60 minutes. The top gel is the PCR product representing the total amount of

transcript and the bottom gel is of the splinted-ligated product (i.e. decapped transcript). There is a

time dependent increase of the splinted-ligated PCR product.

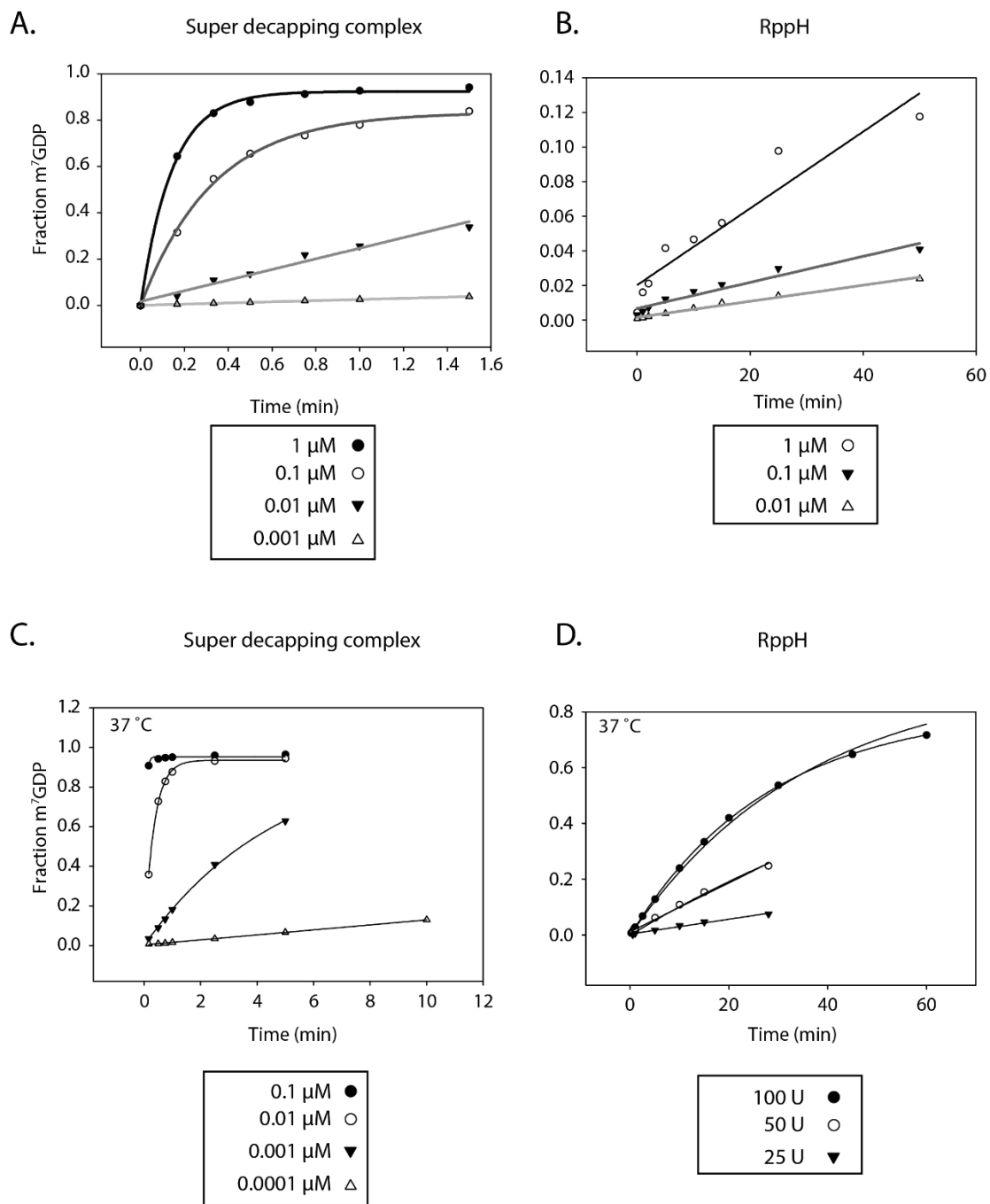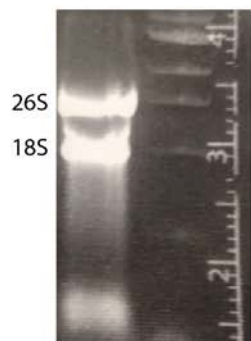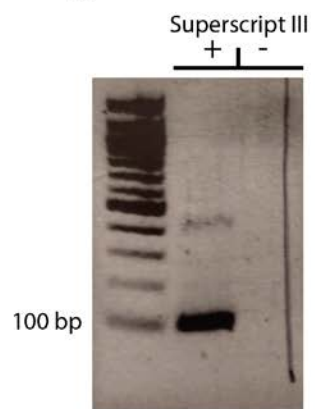Fig S1: Super decapping enzyme is more catalytically efficient than RppH

59

Fig S2: Quality of S. pombe total RNA and gene-specific RT control

A.

26S
18S

15 µg S. pombe total RNA

B.

Superscript III
+   -

100 bp

2 µg input RNA,
Rps23 total mRNA PCR product

SUPPLEMENTARY FIGURE LEGENDS

Fig S1: (A) Plot of fraction m^{7}GDP product versus time for the super decapping enzyme at the concentrations noted in the key. Reactions were performed under single-turnover conditions at 4 °C. (B) Same as in A but with purified RppH. (C) Plot of fraction m^{7}GDP product versus time for the super decapping enzyme under multiple-turnover conditions at 37 °C with *pombe* total RNA. Same as in C but with units of NEB RppH.

Fig S2: (A) Ethidium Bromide stained 1% TAE agarose gel showing the quality of the *pombe* total RNA used in Fig 2C and for the splinted ligation reaction in Fig3. Major ribosomal RNAs are noted. (B) Ethidium Bromide stained 2% TAE agarose gel showing that gene specific RT-PCR of the *Rps23* total mRNA signal is dependent on the inclusion of reverse transcriptase.

CHAPTER 3

A journey in *S. pombe* genetics, insight into Edc3 transcript

regulation, and insect cell expression of full-length fungal Dcp2

I. **Edc3 null strain comparison with WT of transcript dysregulation**

Originally, we had planned on crossing Dcp2 truncations into a Bioneer h+ EDC3::KanMX6 to do similar *in vivo* characterization of the autoinhibitory region as in the Jacobson lab *S. cerevisiae* negative regulatory element paper (He and Jacobson 2015). However, there were several technical difficulties along the way that slowed the experimental progress too much for us to get to the experiments we had wanted to complete. I will detail in this section the progress that had been made, some insights in how we could be more effective going forward, and some preliminary RNA-seq experiments of the *pombe EDC3* null strain.

EDC3 deletion strain shows a mild temperature sensitive phenotype

As a first pass to see if there was any phenotype for the null strain, we performed a serial dilution spotting assay of cells grown in rich, YE5S media, that had been grown to saturation. The cells were diluted to an OD_{595nm} of 0.1 and serially 10-fold diluted. Ten microliters of each dilution was spotted onto a YE5S agar plate that had been uniformly arrayed. The plates were allowed to recover for 2-3 days at either 30 °C or 37 °C for several days before being imaged. The haploid *EDC3* null strain exhibited a reproducible TS phenotype at 37 °C (**Fig 1A**), whereas no appreciable difference was seen under normal growth temperatures. Although, the null strain did exhibit an apparent delay in reaching the exponential growth phase in liquid culture.

Unclear if Dcp2 truncation rescues the TS phenotype of *EDC3* deletion

Previously, the Sprangers group had shown that the removal of the entire *S. pombe* Dcp2 C-terminal extension resulted in a reproducible slow-growth phenotype at 36 °C (Fromm et al. 2012). We were able to reproduce these results in our haploid strain background (**Fig 1A**). As a first pass, we decided

to see if the introduction of a truncated Dcp2 in the *EDC3* null background suppressed the TS phenotypes of the individual strains. The rationale being that *EDC3* deletion results in less active decapping *in vivo*, since removal of the inhibitory region of Dcp2 partially bypasses Edc3 requirement in an *in vitro* decapping assay. As a first pass it appeared that the truncation did suppress the TS; however, the control strains make the interpretation suspect since the reintroduction of the C-terminally extended eGFP fused Dcp2 with a modified 3'UTR also seemed to suppress the TS (**Fig 1B)** At this point, we are unsure if we have bypassed the phenotype. More surgical disruptions of the endogenous locus would have to be constructed in order to get around the odd behavior we noticed with our strains.

Splinted-ligation RT-PCR assay to monitor *in vivo* decapping defects

In published work identifying the *S. pombe* homolog of Edc4 (Pdc1), a group reported monitoring the 5' cap structure of *Rps23* in *pombe* (Wang et al. 2013). They adapted the Coller lab splinted-ligation RT-PCR decapping assay (Blewett et al. 2011) to monitor the abundant *Rps23* ribosomal protein transcript and found that in addition to *PDC1* deletion that *EDC3* deletion also led to a decrease in the amount of capped transcript. We thought that by monitoring this transcript, we might be able to see a bypass of this decapping defect when we excised the C-terminal extension of Dcp2. We generated strains where the endogenous Dcp2 was replaced with Dcp2(1-243)-eGFP fusion in both WT and *EDC3* null haploid backgrounds. We extracted and purified *pombe* total RNA as described and set out to compare the splinted-ligation RT-PCR signal between the various strains.

Unfortunately, we ran into a series of technical difficulties. The main one being that it was difficult to consistently reproduce the results seen by the other group and that the signal for the decapped product was very-weak in all strain backgrounds. We had to go to a high PCR-cycle number and stain for a fairly long time to be able to see a quantifiable product **(Fig 2)**. In fact, in the Coller Methods in Enzymology paper (Blewett et al. 2011), they noted that the signal of the decapped product was only two-fold above

background, which leaves too small of a dynamic range to be able to see a rescue of the decapping defect compared to WT. Deletion of the cytoplasmic exoribonuclease, *XRN1*, resulted in an increase of the decapped product and a larger dynamic range above background. We would have to further delete this gene in our *pombe* strains if we want to be able to determine if Dcp2 truncation can rescue the *EDC3* null decapping phenotype. An attempt was made to remove this gene, but the strains were not tractable in a reasonable timeframe.

RNA-seq reveals larger transcript dysregulation of *EDC3* deletion in *pombe*

Since we already had strains generated, we thought it would be informative if we characterized the transcripts that are affected by *EDC3* deletion in *S. pombe*, since no one had done this bookkeeping. The only comparable data we had was in the *S. cerevisae* system, where *EDC3* deletion only resulted in the stabilization of two transcripts (Badis et al. 2004; Dong et al. 2007). We had reason to believe that *EDC3* in *pombe* might act more generally since the deletion showed an increase in the numbers of p-bodies under rich-media conditions (Wang et al. 2013). An increase in the numbers or size of p-bodies is characteristic of a block in 5' to 3' decay. The classic example is the deletion of the downstream exoribonuclease, Xrn1, which results in an enlargement of p-bodies (Sheth 2003).

As a first pass, we extracted total RNA from mid-log phase WT and *EDC3* null h+ haploid cells and enriched for mRNAs by oligodT selection. While ribo-minus selection would be more ideal for monitoring a decapping defect stabilization of transcripts, since deadenylated capped RNAs would accumulate and these could be loss with oligodT based isolation. However, our collaborators did not have ribo-minus enrichment available. Regardless, this dataset would at the very least give us insight on whether Edc3 in *pombe* might have a different role. From our first-pass RNA-seq data, it is apparent that something different is happening in fission yeast. Hundreds of transcripts are dysregulated in *pombe* compared to

only two in *S. cerevisiae* (**Fig 3**). Furthermore, none of the affected transcripts are homologs to the *S. cerevisiae* transcripts.

Interestingly, the large number of transcripts that are either increased or decreased in abundance upon *EDC3* deletion are associated with NAD/PH metabolism or binding. This is of note, since the dimerization domain of Edc3 is a YjeF N enzymatic fold that in *S. cerevisiae* had been shown *in vitro* to have an enzymatic effect on NAD (Walters et al. 2014). If we were going to determine the transcripts that are attributed to Edc3's role in decapping, then we would have to be more surgical in our disruption of its function. Targeted mutations to the LSm or the YjeF N domains could allow us to separate the transcripts that are affected by the activation of Dcp2 or the putative enzymatic role of its catalytic dimerization domain. Other transcripts of note are listed in (**Table 1)** and potentially implicate Edc3 in RNAi mediated gene-silencing in *pombe*. Further work is required to follow up on these novel results.

II. Mutation aromatic/aliphatic series of Dcp2

In work by Christina Fitzsimmons (a rotation student) and guided by my design, we tested a mutational series of the gatekeeper W43 residue to see if we could better determine if Dcp2 was undergoing a linear open to close (or vice-versa) transition. The expected trend was that increasing hydrophobic surface area (HSA) of the sidechain would result in an activity more similar to WT. W43 was mutated to Tyr, Phe, Ile, and Ala and the activities were compared in an *in vitro* decapping assay. The predicted trend held up as expected (**Fig 4A)**. Additionally, if Dcp2 were undergoing some two-state transition then we expected to see collinear shifts from the W43A to W43 WT with the other mutations falling in between with the more HAS sidechains closer in Chemical shift perturbations to WT. For the most part that is what we saw (**Fig 4B**).

Finally, as a test for our plasmid shuffling Dcp2 strain that we had constructed, we wanted to see what substitutions at W43 were viable. Dcp2 is an essential gene in *S. pombe*, so we replaced the endogenous copy with a plasmid borne Dcp2 that could be counter-selected by 5-FOA (**Fig 4C**). A series of W43 substitutions on *Leu2* containing plasmids were introduced and maintained on -Leu media, which allowed for the WT plasmid containing *Ura4* to be lost if the mutant Dcp2 is functional. Only cells that have lost the original WT plasmid will grow on 5-FOA media; any other cells with a functional *Ura4* will be killed. Under these conditions, we see that only aromatic substitutions of W43 are viable **(Fig 4D)**. Presumably, this is due to W43 being important for binding the m^7G cap during catalysis by stacking with the base (Mugridge et al. 2016; Charenton et al. 2016; Wurm et al. 2017); something that is only possible by aromatic amino acids.

III. Insect-cell (SF9) expression of full-length *pombe* Dcp2

Prior to our successful expression, purification and biochemical characterization of C-terminally extended Dcp2 from *E. coli*, we had initially done test expression and purifications of full-length *pombe* Dcp2 in SF9 cells. Below is a synopsis of the plasmids we constructed to generate the bacmids, the amplification of the viruses, and the results of the test expressions.

Construction of the pFastBac for Dcp2-eGFP and Edc3-eGFP

The cDNA of *pombe* Dcp2 or Edc3 were cloned into the MCS of a standard pFastBac plasmid that included a Tev-cleavable C-terminal eGFP-StepII tag. The eGFP allows for monitoring the infection and generation of virus and protein during viral passage and overexpression respectively. Monitoring of cellular eGFP revealed that both proteins appeared to be stably expressed in SF9 cells (**Fig 5A)**; however, upon purification of the constructs it was apparent that full-length Dcp2 was prone to degradation. Edc3

was stably expressed and mostly intact and folded protein was recovered (**Fig 5B**). A different strategy to purify Dcp2 was needed.

Coinfection of Dcp2 and Edc3 in SF9 rescues full-length Dcp2

Since Edc3 binds multiple HLMs on the disordered tail of Dcp2 (Fromm et al. 2014), we hypothesized that coinfection of the Edc3-eGFP containing baculovirus would stabilize this degradation prone extension. SF9 cells were infected with the same amount of total virus but instead of a single virus it consisted of a 50/50 mix of both Edc3 and Dcp2 viruses. After 48 hours of infection, we imaged a small sample of cells to check for eGFP. Interestingly, we noticed a difference in the expression pattern of the dual infection (**Fig 5C**). Unlike the individual infections, here we see what look like phase-separated puncta in the SF9 cells. Dcp2 and Edc3 are known to form liquid-liquid demixed phases *in vitro* (Fromm et al. 2014), so we presume that we are seeing a similar phenomenon in the SF9 cells. Consistent with this we also saw a similar expression pattern in our pFastBac-dual bacmid containing Dcp2-eGFP and untagged Edc3. Purification of the coinfected SF9 cells resulted in a recovery of full-length Dcp2 (**Fig 5D**). However, we do not know the stoichiometry of this complex nor were we able to separate Dcp2 from Edc3. Therefore, this expression system is not ideal for being able to biochemically characterize the activity of full-length Dcp2 or the effect Edc3 has on Dcp2's activity *in vitro*.

Table 1: Selection of transcripts that have altered abundance in *EDC3 h+* haploid deletion

NAD/PH and metabolic related genes

Gene standard name	Log$_2$ Fold Change	Systematic ID	Function
unassigned	-1.868543026	SPBC1271.07c	Acyl-CoA- N-acyltransferase
unassigned	-1.764970377	SPAC5H10.10	NADPH dehydrogenase
DEA2	-1.506954453	SPBC1198.02	adenine deaminase
unassigned	-1.345494836	SPAC11D3.02c	Acyl-CoA- N-acyltransferase
ADH4	-1.33990939	SPAC5H10.06c	alcohol dehydrogenase
unassigned	-1.256349751	SPAC8E11.10	sorbase reductase
GDH2	-1.064774478	SPCC132.04c	NAD-dependent glutamate dehydrogenase
unassigned	-1.015606823	SPAPB24D3.08c	NADP-dependent oxidoreductase
GPD3	1.030404682	SPBC354.12	glyceraldehyde 3-phosphate dehydrogenase
GAL10	1.27250852	SPBPB2B2.12c	UDP-glucose 4-epimerase/aldose 1-epimerase
unassigned	1.626792788	SPACUNK4.17	NAD binding dehydrogenase family protein

RNAi mediated gene silencing and mRNA decay

Gene standard name	Log$_2$ Fold Change	Systematic ID	Function
AES1	-1.389916662	SPBPB21E7.07	Enhancer or RNA-mediated gene silencing
CNP3	-1.33501384	SPBC1861.01c	protein required for centromere silencing
PUF5	-1.070275242	SPAC4G8.03c	nuclear transcribed mRNA process, deadenylation-dependent decay

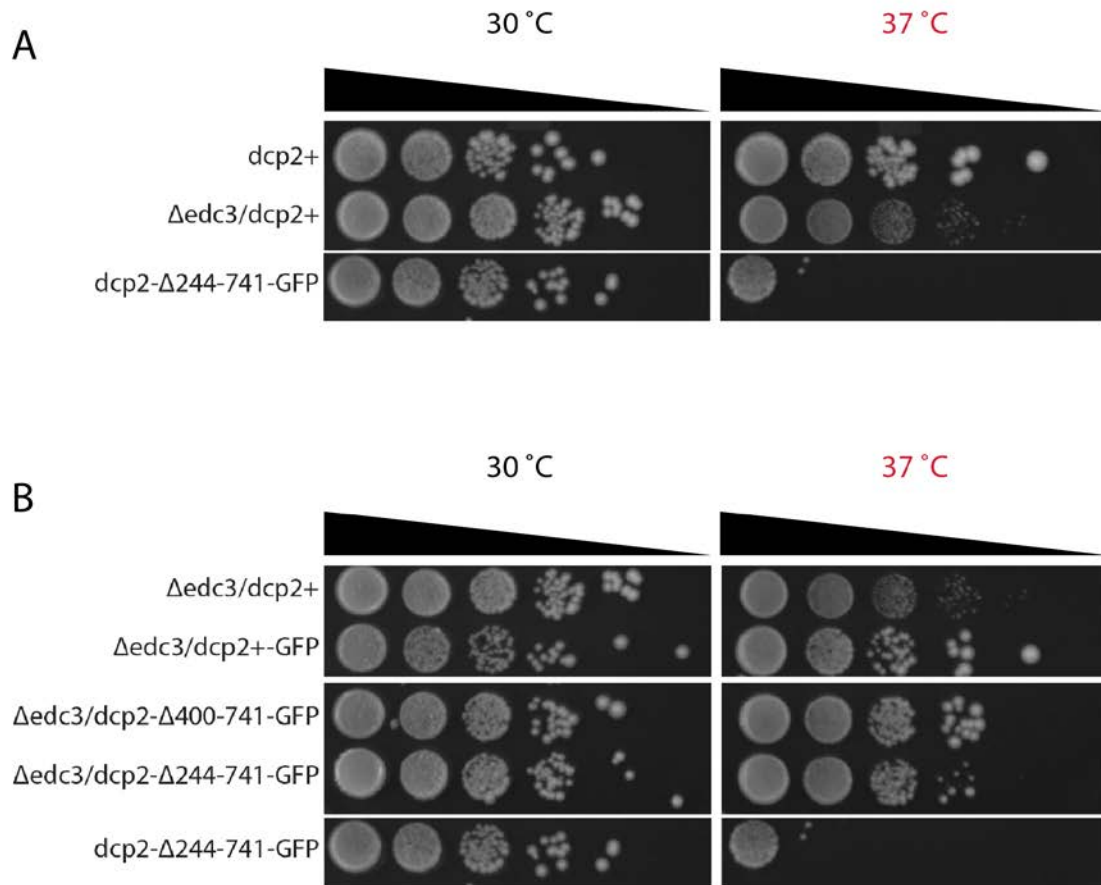Fig. 1: EDC3 null strain and Dcp2 truncation have TS phenotype

A

30 ˚C   37 ˚C

dcp2+
Δedc3/dcp2+
dcp2-Δ244-741-GFP

B

30 ˚C   37 ˚C

Δedc3/dcp2+
Δedc3/dcp2+-GFP
Δedc3/dcp2-Δ400-741-GFP
Δedc3/dcp2-Δ244-741-GFP
dcp2-Δ244-741-GFP

Fig. 2: Endogenous decapped Rps23 mRNA is difficult to detect by SL-RTPCR
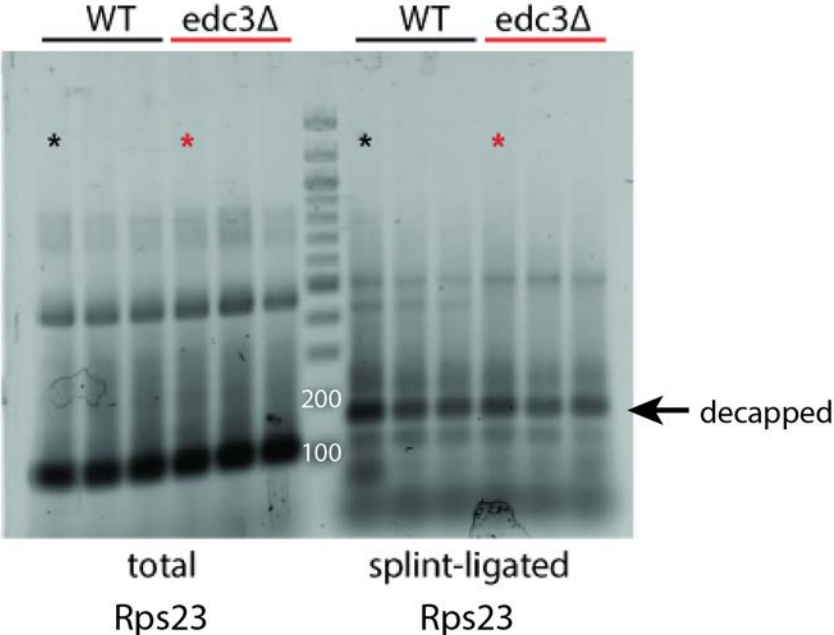

total
Rps23

splint-ligated
Rps23

Fig. 3 Hundreds of transcript levels are altered in EDC3 deletion
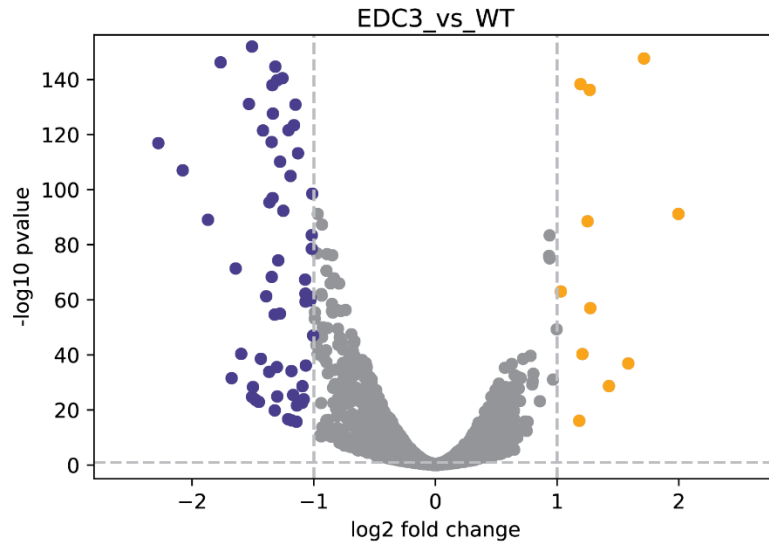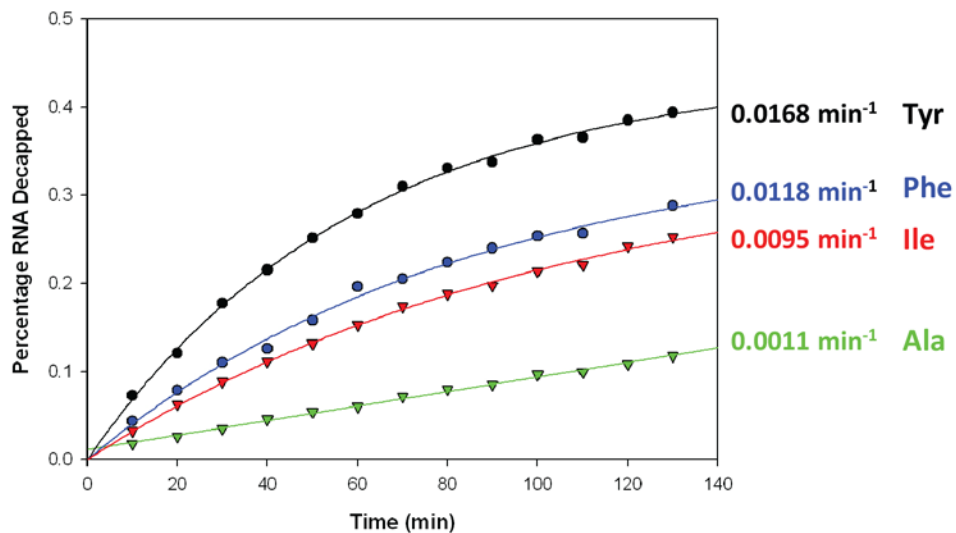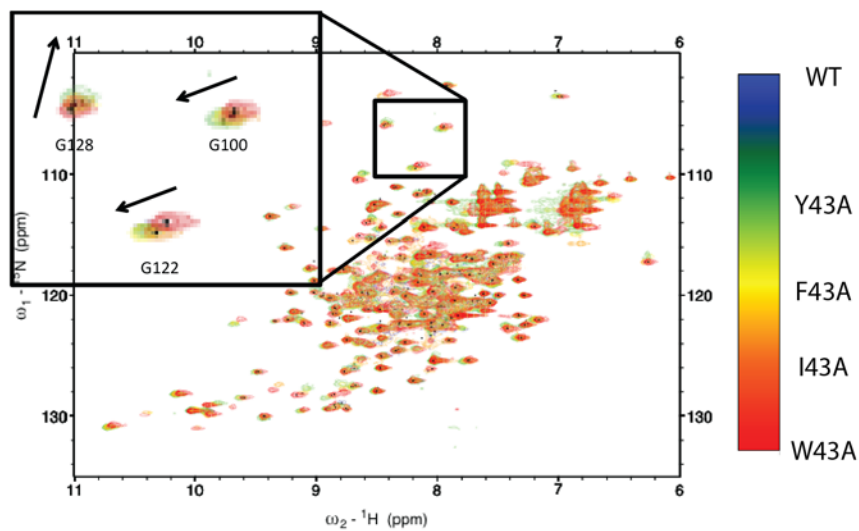

EDC3_vs_WT

Fig. 4: Activity, dynamics, and viability of W43 mutants as a function of side-chain hydrophobic surface area

A

B

C
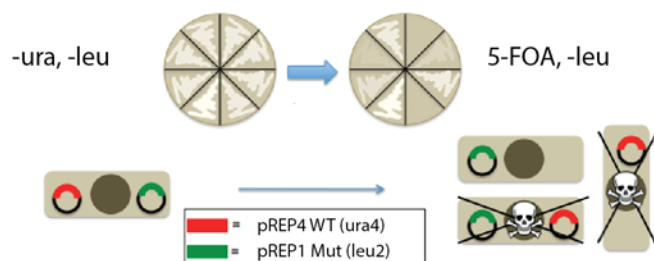
D

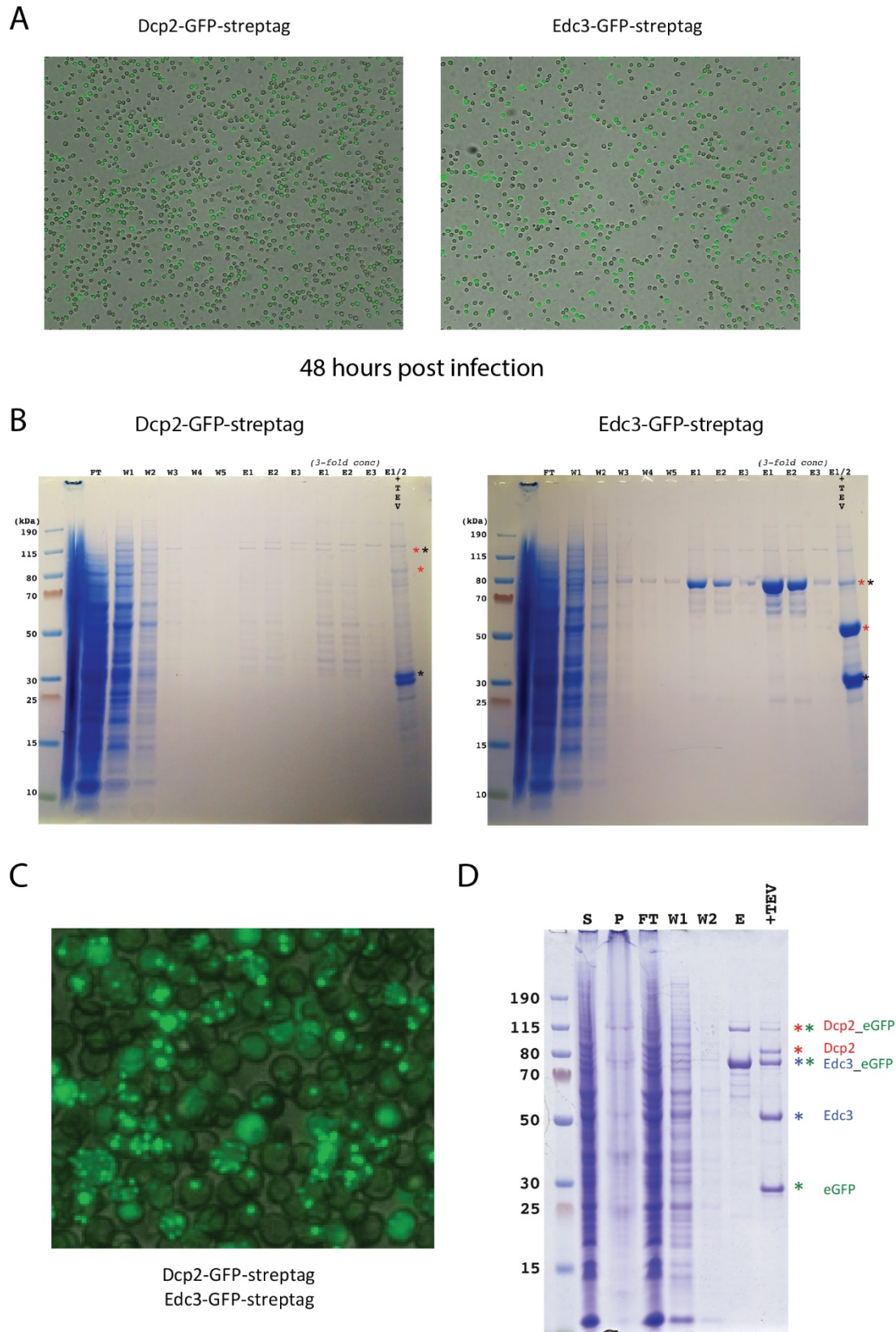# Fig. 5: Full-length spDcp2 can be stably coexpressed with Edc3 in SF9 cells

A

Dcp2-GFP-streptag

Edc3-GFP-streptag

48 hours post infection

B

Dcp2-GFP-streptag

Edc3-GFP-streptag

C

Dcp2-GFP-streptag
Edc3-GFP-streptag

D

FIGURE LEGENDS

Fig 1: (A) 10-fold Serial dilution spotting of 0.1 OD_{595nm} of cells from the strains as noted grown at 30 °C and 37 °C. (B) Same as in A, but showing suppression of the Dcp2 C-terminal truncation TS phenotype.

Fig 2: Sybr Gold stained 2% TAE agarose gel of the splinted-ligation RT-PCR reaction monitoring both total *Rps23* or the ligated, decapped *Rps23* for wild-type and *EDC3* null strains.

Fig 3: Volcano plot of the $\log_2$ fold change of transcripts in the *EDC3* null compared to wild-type. Colored dots represent transcripts that have significantly changed; blue is decreased and yellow is increased.

Fig 4: (A) Fraction of m^7GDP versus time for a mutational series of W43 in Dcp2 (1-243). (B) N^{15}-HSQC of the same mutants monitoring peaks that display co-linear shifts depending on the severity of the mutation. (C) Schematic of the plasmid-shuffling assay for essential genes. (D) Viability of the W43 mutants in *S. pombe* with control K129 catalytic site mutant.

Fig 5: (A) Visualization of GFP fluorescence in infected SF9 cells for the expression test of full-length *pombe* Dcp2 and Edc3. (B) SDS-PAGE of Strep-tactin purification of Dcp2 and Edc3 followed by Tev cleavage; red star is protein of interest and black star is eGFP. (C) Coinfection of Dcp2 and Edc3 form puncta as judged by GFP fluorescence. (D) SDS-PAGE of purified Dcp2:Edc3 complex from the coinfection showing un-cleaved and cleaved samples.

OUTRO

During the course of studying the *S. pombe* C-terminally extension of Dcp2, I have stumbled upon some novel results. While other labs have shown *in vivo* or hypothesized that the decapping complex can be negatively regulated, no one to our knowledge has shown that Dcp2 is autoinhibited. Furthermore, this autoinhibited complex recapitulates a more relevant system for us to understand the functional role of Edc3 in *S. pombe* as a possible more general decapping activator. However, more detailed studies *in vivo* need to be realized before we can truly understand the entirety of the Edc3 functional role in fission yeast. At the most basic level, Edc3 regulation of transcripts appears to be more complicated in *S. pombe* based upon the greater number of transcripts that are dysregulated. It will be of great interest to see a finer study dissecting the functional roles of the LSm and YjeF N domains of Edc3. The autoinhibited decapping complex adds another layer to the regulation of 5' mediated decay, and opens up the possibility that we might be able to better explain *in vivo* phenomenon that have been difficult to resolve in previous *in vitro* systems.

I was unable to realize my goal of seeing if the *EDC3* null decapping defect was a result of the autoinhibition. The expectation was that introduction of a Dcp2 lacking IM1 and IM2 would suppress this defect. Despite this, my foray in *S. pombe* genetic was informative and may shed some light in the future in regard to the transcripts that are regulated by Edc3. My attempts to replicate a published splinted-ligation result were enlightening in helping me understand how small details in protocols can have large impacts in the final product. It was especially enlightening in how often these details are not carefully noted in protocols. The splinted-ligation assay troubleshooting led to our demonstration of using of a constitutively-active decapping complex as a replacement for TAP. This reagent could make for a better replacement than RppH for sequencing methods that rely on removal of the cap for a ligateable cloning handle. Overall this has been a rewarding, but trying Doctoral dissertation.

REFERENCES

Ahmed I, Buchert R, Zhou M, Jiao X, Mittal K, Sheikh TI, Scheller U, Vasli N, Rafiq MA, Qasim Brohi M, et
al. 2014. Mutations in DCPS and EDC3 in autosomal recessive intellectual disability indicate a
crucial role for mRNA decapping in neurodevelopment. *Hum Mol Genet* **24**: 3172–3180.

Arribas-Layton M, Wu D, Lykke-Andersen J, Song H. 2013. Structural and functional control of the
eukaryotic mRNA decapping machinery. *Biochim Biophys Acta - Gene Regul Mech* **1829**: 580–589.
http://dx.doi.org/10.1016/j.bbagrm.2012.12.006.

Badis G, Saveanu C, Fromont-Racine M, Jacquier A. 2004. Targeted mRNA degradation by deadenylation-
independent decapping. *Mol Cell* **15**: 5–15.

Bauer NC, Doetsch PW, Corbett AH. 2015. Mechanisms Regulating Protein Localization. *Traffic* **16**: 1039–
1061.

Blewett N, Coller J, Goldstrohm A. 2011. A quantitative assay for measuring mRNA decapping by splinted
ligation reverse transcription polymerase chain reaction: qSL-RT-PCR. *Rna* **17**: 535–543.
http://rnajournal.cshlp.org/cgi/doi/10.1261/rna.2436411.

Borja MS, Piotukh K, Freund C, Gross JD. 2011. Dcp1 links coactivators of mRNA decapping to Dcp2 by
proline recognition. *Rna* **17**: 278–290. http://rnajournal.cshlp.org/cgi/doi/10.1261/rna.2382011.

Brown CY, Mize GJ, Pineda M, George DL, Morris DR. 1999. Role of two upstream open reading frames in
the translational control of oncogene mdm2. *Oncogene* **18**: 5631–5637.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_u
ids=10523842%5Cnhttp://www.nature.com/onc/journal/v18/n41/pdf/1202949a.pdf.

Calvo SE, Pagliarini DJ, Mootha VK. 2009. Upstream open reading frames cause widespread reduction of

protein expression and are polymorphic among humans. *Proc Natl Acad Sci* **106**: 7507–7512.

http://www.pnas.org/cgi/doi/10.1073/pnas.0810916106.

Chang JH, Jiao X, Chiba K, Oh C, Martin CE, Kiledjian M, Tong L. 2012. Dxo1 is a new type of eukaryotic

enzyme with both decapping and 5'-3' exoribonuclease activity. *Nat Struct Mol Biol* **19**: 1011–1017.

http://www.nature.com/doifinder/10.1038/nsmb.2381.

Charenton C, Taverniti V, Gaudon-Plesse C, Back R, Séraphin B, Graille M. 2016. Structure of the active

form of Dcp1–Dcp2 decapping enzyme bound to m7GDP and its Edc3 activator. *Nat Struct Mol Biol*

23: 982–986. http://www.nature.com/doifinder/10.1038/nsmb.3300.

Cho H, Park OH, Park J, Ryu I, Kim J, Ko J, Kim YK. 2015. Glucocorticoid receptor interacts with PNRC2 in a

ligand-dependent manner to recruit UPF1 for rapid mRNA degradation. *Proc Natl Acad Sci* **112**:

E1540–E1549. http://www.pnas.org/lookup/doi/10.1073/pnas.1409612112.

Chowdhury A, Tharun S. 2008. lsm1 mutations impairing the ability of the Lsm1p-7p-Pat1p complex to

preferentially bind to oligoadenylated RNA affect mRNA decay in vivo. *Rna* **14**: 2149–2158.

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2553750&tool=pmcentrez&rendertyp

e=abstract%5Cnhttp://www.rnajournal.org/cgi/doi/10.1261/rna.1094208.

Coller J, Parker R. 2005. General translational repression by activators of mRNA decapping. *Cell* **122**:

875–886.

Davey NE, Cyert MS, Moses AM. 2015. Short linear motifs – ex nihilo evolution of protein regulation. *Cell*

Commun Signal **13**: 43. http://www.biosignaling.com/content/13/1/43.

Deana A, Celesnik H, Belasco JG. 2008. The bacterial enzyme RppH triggers messenger RNA degradation

by 5' pyrophosphate removal. *Nature* **451**: 355–358.

http://www.nature.com/doifinder/10.1038/nature06475.

Decker CJ, Parker R. 2012. P-Bodies and Stress Granules : Possible Roles in the Control of Translation and

mRNA Degradation.

Decker CJ, Teixeira D, Parker R. 2007. Edc3p and a glutamine/asparagine-rich domain of Lsm4p function

in processing body assembly in Saccharomyces cerevisiae. *J Cell Biol* **179**: 437–449.

Dieudonné F-X, O'Connor PBF, Gubler-Jaquier P, Yasrebi H, Conne B, Nikolaev S, Antonarakis S, Baranov

P V., Curran J. 2015. The effect of heterogeneous Transcription Start Sites (TSS) on the translatome:

implications for the mammalian cellular phenotype. *BMC Genomics* **16**: 986.

http://www.biomedcentral.com/1471-2164/16/986.

Dong S, Li C, Zenklusen D, Singer RH, Jacobson A, He F. 2007. YRA1 Autoregulation Requires Nuclear

Export and Cytoplasmic Edc3p-Mediated Degradation of Its Pre-mRNA. *Mol Cell* **25**: 559–573.

Dosztányi Z, Csizmok V, Tompa P, Simon I. 2005. IUPred: Web server for the prediction of intrinsically

unstructured regions of proteins based on estimated energy content. *Bioinformatics* **21**: 3433–

3434.

Dunckley T, Parker R. 1999. The DCP2 protein is required for mRNA decapping in Saccharomyces

cerevisiae and contains a functional MutT motif. *EMBO J* **18**: 5411–5422.

Eulalio A, Rehwinkel J, Stricker M, Huntzinger E, Yang SF, Doerks T, Dorner S, Bork P, Boutros M,

Izaurralde E. 2007. Target-specific requirements for enhancers of decapping in miRNA-mediated

gene silencing. *Genes Dev* **21**: 2558–2570.

Farcas AM, Blackledge NP, Sudbery I, Long HK, McGouran JF, Rose NR, Lee S, Sims D, Cerase A, Sheahan

TW, et al. 2012. KDM2B links the polycomb repressive complex 1 (PRC1) to recognition of CpG

islands. *Elife* **2012**: 1–26.

Floor SN, Borja MS, Gross JD. 2012. Interdomain dynamics and coactivation of the mRNA decapping

enzyme Dcp2 are mediated by a gatekeeper tryptophan. *Proc Natl Acad Sci* **109**: 2872–2877. http://www.pnas.org/cgi/doi/10.1073/pnas.1113620109.

Floor SN, Jones BN, Hernandez G a, Gross JD. 2010. A split active site couples cap recognition by Dcp2 to activation. *Nat Struct Mol Biol* **17**: 1096–1101. http://dx.doi.org/10.1038/nsmb.1879.

Fromm SA, Kamenz J, N??ldeke ER, Neu A, Zocher G, Sprangers R. 2014. InVitro reconstitution of a cellular phase-transition process that involves the mRNA decapping machinery. *Angew Chemie - Int Ed* **53**: 7354–7359.

Fromm SA, Truffault V, Kamenz J, Braun JE, Hoffmann NA, Izaurralde E, Sprangers R. 2012. The structural basis of Edc3- and Scd6-mediated activation of the Dcp1:Dcp2 mRNA decapping complex. *EMBO J* **31**: 279–290. http://emboj.embopress.org/cgi/doi/10.1038/emboj.2011.408.

Garneau NL, Wilusz J, Wilusz CJ. 2007. The highways and byways of mRNA decay. *Nat Rev Mol Cell Biol* **8**: 113–126. http://www.nature.com/doifinder/10.1038/nrm2104.

Gu W, Lee HC, Chaves D, Youngman EM, Pazour GJ, Conte D, Mello CC. 2012. CapSeq and CIP-TAP identify pol ii start sites and reveal capped small RNAs as C. elegans piRNA precursors. *Cell* **151**: 1488–1500. http://dx.doi.org/10.1016/j.cell.2012.11.023.

Harigaya Y, Jones BN, Muhlrad D, Gross JD, Parker R. 2010. Identification and analysis of the interaction between Edc3 and Dcp2 in Saccharomyces cerevisiae. *Mol Cell Biol* **30**: 1446–56. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2832485&tool=pmcentrez&rendertype=abstract.

He F, Jacobson A. 2015. Control of mRNA decapping by positive and negative regulatory elements in the Dcp2 C-terminal domain. *RNA* 1633–1647. http://www.ncbi.nlm.nih.gov/pubmed/26184073.

Henzler-Wildman K, Kern D. 2007. Dynamic personalities of proteins. *Nature* **450**: 964–972.

http://www.nature.com/doifinder/10.1038/nature06522.

Izaurralde E, Lewis J, McGuigan C, Jankowska M, Darzynkiewicz E, Mattaj IW. 1994. A nuclear cap

binding protein complex involved in pre-mRNA splicing. *Cell* **78**: 657–68.

http://www.ncbi.nlm.nih.gov/pubmed/8069914.

Jin X, Turcott E, Englehardt S, Mize GJ, Morris DR. 2003. The two upstream open reading frames of

oncogene mdm2 have different translational regulatory properties. *J Biol Chem* **278**: 25716–25721.

Jonas S, Izaurralde E. 2013. The role of disordered protein regions in the assembly of decapping

complexes and RNP granules. *Genes Dev* **27**: 2628–2641.

Jones BN, Quang-Dang DU, Oku Y, Gross JD. 2008. *Chapter 2 A Kinetic Assay to Monitor RNA Decapping

Under Single- Turnover Conditions*. 1st ed. Elsevier Inc. http://dx.doi.org/10.1016/S0076-

6879(08)02602-5.

Kapranov P. 2009. From transcription start site to cell biology. *Genome Biol* **10**: 217.

http://genomebiology.biomedcentral.com/articles/10.1186/gb-2009-10-4-217.

Konarska MM, Padgett RA, Sharp PA. 1984. Recognition of cap structure in splicing in vitro of mRNA

precursors. *Cell* **38**: 731–736.

Lai T, Cho H, Liu Z, Bowler MW, Piao S, Parker R, Kim YK, Song H. 2012. Structural basis of the PNRC2-

mediated link between mRNA surveillance and decapping. *Structure* **20**: 2025–2037.

http://dx.doi.org/10.1016/j.str.2012.09.009.

Latysheva NS, Flock T, Weatheritt RJ, Chavali S, Babu MM. 2015. How do disordered regions achieve

comparable functions to structured domains? *Protein Sci* **24**: 909–922.

Ling SHM, Decker CJ, Walsh M a, She M, Parker R, Song H. 2008. Crystal structure of human Edc3 and its

functional implications. *Mol Cell Biol* **28**: 5965–76.

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2547010&tool=pmcentrez&rendertyp

e=abstract%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/18678652%5Cnhttp://www.pubmedcentra

l.nih.gov/articlerender.fcgi?artid=PMC2547010.

Lykke-Andersen S, Jensen TH. 2015. Nonsense-mediated mRNA decay: an intricate machinery that

shapes transcriptomes. *Nat Rev Mol Cell Biol* **16**: 665–677.

http://www.nature.com/doifinder/10.1038/nrm4063.

Mauer J, Luo X, Blanjoie A, Jiao X, Grozhik A V., Patil DP, Linder B, Pickering BF, Vasseur J-J, Chen Q, et al.

2016. Reversible methylation of m6Am in the 5' cap controls mRNA stability. *Nature* **541**: 371–375.

http://www.nature.com/doifinder/10.1038/nature21022.

Molleston JM, Cherry S. 2017. Attacked from all sides: RNA decay in antiviral defense. *Viruses* **9**.

Moore MJ. 2005. From Birth to Death: The Complex Lives of Eukaryotic mRNAs. *Science (80-)* **309**:

1514–1518. http://www.sciencemag.org/cgi/doi/10.1126/science.1111443.

Mugridge JS, Ziemniak M, Jemielity J, Gross JD. 2016. Structural basis of mRNA-cap recognition by Dcp1–

Dcp2. *Nat Struct Mol Biol* **23**: 987–994. http://www.nature.com/doifinder/10.1038/nsmb.3301.

Nagarajan VK, Jones CI, Newbury SF, Green PJ. 2013. XRN 5'→3' exoribonucleases: Structure,

mechanisms and functions. *Biochim Biophys Acta - Gene Regul Mech* **1829**: 590–603.

http://dx.doi.org/10.1016/j.bbagrm.2013.03.005.

Narlikar GJ, Fan HY, Kingston RE. 2002. Cooperation between complexes that regulate chromatin

structure and transcription. *Cell* **108**: 475–487.

Newbury S, Woollard A. 2004. The 5–3 exoribonuclease. 59–65.

Nissan T, Rajyaguru P, She M, Song H, Parker R. 2010. Decapping Activators in Saccharomyces cerevisiae Act by Multiple Mechanisms. *Mol Cell* **39**: 773–783. http://dx.doi.org/10.1016/j.molcel.2010.08.025.

Pelechano V, Wei W, Steinmetz LM. 2013. Extensive transcriptional heterogeneity revealed by isoform profiling. *Nature* **497**: 127–131. http://www.nature.com/doifinder/10.1038/nature12121.

Radhakrishnan A, Chen YH, Martin S, Alhusaini N, Green R, Coller J. 2016. The DEAD-Box Protein Dhh1p Couples mRNA Decay and Translation by Monitoring Codon Optimality. *Cell* **167**: 122–132.e9. http://dx.doi.org/10.1016/j.cell.2016.08.053.

Ravid T, Hochstrasser M. 2008. Diversity of degradation signals in the ubiquitin–proteasome system. *Nat Rev Mol Cell Biol* **9**: 679–689. http://www.nature.com/doifinder/10.1038/nrm2468.

Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**: 320–324.

Rojas-Duran MF, Gilbert W V. 2012. Alternative transcription start site selection leads to large differences in translation activity in yeast. *Rna* **18**: 2299–2305. http://rnajournal.cshlp.org/cgi/doi/10.1261/rna.035865.112.

Rudnizky S, Malik O, Bavly A, Pnueli L, Melamed P, Kaplan A. 2017. Nucleosome mobility and the regulation of gene expression: Insights from single-molecule studies. *Protein Sci* **0**: 1–12.

Schoenberg DR, Maquat LE. 2012. Regulation of cytoplasmic mRNA decay. *Nat Rev Genet* **13**: 448–448. http://www.nature.com/doifinder/10.1038/nrg3254.

Schütz S, Nöldeke ER, Sprangers R. 2017. A synergistic network of interactions promotes the formation of in vitro processing bodies and protects mRNA against decapping. *Nucleic Acids Res* 1–12. http://www.ncbi.nlm.nih.gov/pubmed/28472520%0Ahttps://academic.oup.com/nar/article-

lookup/doi/10.1093/nar/gkx353.

SGrudzien-Nogalska E, Kiledjian M. 2017. New insights into decapping enzymes and selective mRNA decay. *Wiley Interdiscip Rev RNA* **8**: 1–11.

She M, Decker CJ, Svergun DI, Round A, Chen N, Muhlrad D, Parker R, Song H. 2008. Structural Basis of Dcp2 Recognition and Activation by Dcp1. *Mol Cell* **29**: 337–349.

Sheth U. 2003. Decapping and Decay of Messenger RNA Occur in Cytoplasmic Processing Bodies. *Science (80-)* **300**: 805–808. http://www.sciencemag.org/cgi/doi/10.1126/science.1082320.

Song MG, Li Y, Kiledjian M. 2010. Multiple mRNA Decapping Enzymes in Mammalian Cells. *Mol Cell* **40**: 423–432. http://dx.doi.org/10.1016/j.molcel.2010.10.010.

Tompa P. 2003. Intrinsically unstructured proteins evolve by repeat expansion. *BioEssays* **25**: 847–855.

Topisirovic I, Svitkin Y V., Sonenberg N, Shatkin AJ. 2011. Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip Rev RNA* **2**: 277–298.

Tsompana M, Buck MJ. 2014. Chromatin accessibility: a window into the genome. *Epigenetics Chromatin* **7**: 33. http://www.ncbi.nlm.nih.gov/pubmed/25473421%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4253006.

Valkov E, Muthukumar S, Chang C-T, Jonas S, Weichenrieder O, Izaurralde E. 2016. Structure of the Dcp2–Dcp1 mRNA-decapping complex in the activated conformation. *Nat Struct Mol Biol* **23**: 574–579. http://www.nature.com/doifinder/10.1038/nsmb.3232.

Visa N, Izaurralde E, Ferreira J, Daneholt B, Mattaj IW. 1996. A nuclear cap-binding complex binds Balbiani ring pre-mRNA cotranscriptionally and accompanies the ribonucleoprotein particle during

nuclear export. *J Cell Biol* **133**: 5–14.

Walters RW, Shumilin IA, Yoon J-H, Minor W, Parker R. 2014. Edc3 function in yeast and mammals is modulated by interaction with NAD-related compounds. *G3 (Bethesda)* **4**: 613–22. http://www.ncbi.nlm.nih.gov/pubmed/24504254%5Cnhttp://www.pubmedcentral.nih.gov/articler ender.fcgi?artid=PMC4059234%5Cnhttp://g3journal.org/cgi/doi/10.1534/g3.114.010470.

Wang C-Y, Chen W-L, Wang S-W. 2013. Pdc1 Functions in the Assembly of P Bodies in Schizosaccharomyces pombe. *Mol Cell Biol* **33**: 1244–1253. http://mcb.asm.org/cgi/doi/10.1128/MCB.01583-12.

Wang Z, Jiao X, Carr-Schmid A, Kiledjian M. 2002. The hDcp2 protein is a mammalian mRNA decapping enzyme. *Proc Natl Acad Sci U S A* **99**: 12663–8. http://www.ncbi.nlm.nih.gov/pubmed/12218187%5Cnhttp://www.pubmedcentral.nih.gov/articler ender.fcgi?artid=PMC130517.

Wurm JANP, Overbeck JAN, Sprangers R. 2016. The S . pombe mRNA decapping complex recruits cofactors and an Edc1-like activator through a single dynamic surface. *Rna* 1–13. http://www.ncbi.nlm.nih.gov/pubmed/27354705.

Wurm JP, Holdermann I, Overbeck JH, Mayer PHO, Sprangers R. 2017. Changes in conformational equilibria regulate the activity of the Dcp2 decapping enzyme. *Proc Natl Acad Sci* 201704496. http://www.pnas.org/lookup/doi/10.1073/pnas.1704496114.

Xiang S, Cooper-Morgan A, Jiao X, Kiledjian M, Manley JL, Tong L. 2009. Structure and function of the 5'→3' exoribonuclease Rat1 and its activating partner Rai1. *Nature* **458**: 784–788. http://www.nature.com/doifinder/10.1038/nature07731.
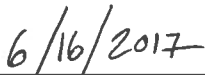
Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

_____   6/16/2017
Author Signature   Date