MCB Accepts, published online ahead of print on 13 December 2010
Mol. Cell. Biol. doi:10.1128/MCB.00799-10
Copyright © 2010, American Society for Microbiology and/or the Listed Authors/Institutions. All Rights Reserved.

MCB Accepts published online ahead of print

# Basic-Leucine Zipper Protein Cnc-C is a Substrate and Transcriptional Regulator of the *Drosophila* 26S Proteasome

Kristian Björk Grimberg#, Anne Beskow#, Daniel Lundin, Monica M. Davis, Patrick Young*
#Co-first Authors

Department of Molecular Biology and Functional Genomics, Stockholm University, Stockholm S-10691, Sweden

Running Title: *Drosophila* proteasome transcription factors

Keywords: proteasome, transcription factors, RNAi screen, Cnc, basic-leucine zipper proteins

---

This work was supported by grants from the Swedish Research Council, and Swedish Cancer Society

* Corresponding Author Tel: + 46-(0)8-52487387

E-mail gly@me.com

MCB Accepts published online ahead of print

**Abstract**

While the 26S proteasome is a key proteolytic complex, little is known about how proteasome levels are maintained in higher eukaryotic cells. Here we describe an RNA interference (RNAi) screen in *Drosophila* that was used to identify transcription factors that may play a role in maintaining levels of the 26S proteasome. We used an RNAi library against 993 *Drosophila* transcription factor genes to identify genes whose suppression in Schneider 2 cells stabilized a ubiquitin-GFP reporter protein. This screen identified Cnc (cap-n-collar/basic-region leucine zipper) as a candidate transcriptional regulator of proteasome component expression. In fact, 20S proteasome activity was reduced in cells depleted of *cnc*. Immunoblots against proteasome components revealed a general decline in both 19S regulatory complex and 20S proteasome subunits after RNAi depletion of this transcription factor. Transcript specific silencing revealed that the longest of the seven transcripts for the *cnc* gene, *cnc-C*, was needed for proteasome and p97 ATPase production. Quantitative RT-PCR confirmed the role of Cnc-C in activation of transcription of genes encoding proteasome components. Expression of a V5His tagged form of Cnc-C revealed that the transcription factor is itself a proteasome substrate that is stabilized when the proteasome is inhibited. We propose that this single *cnc* gene in *Drosophila* resembles the ancestral gene family of mammalian nuclear factor erythroid derived 2-related transcription factors that are essential in regulating oxidative stress and proteolysis.

MCB Accepts published online ahead of print

## Introduction

The ubiquitin/proteasome system (UPS) degrades intracellular proteins and is essential for regulating a wide range of cellular pathways. The UPS plays a critical role in the regulated degradation of proteins involved in tumor development and cell cycle control. Proteins destined for proteasomal degradation are modified by conjugation of ubiquitin moieties through the concerted action of E1, E2, and E3 enzymes. Repeated rounds of conjugation lead to the formation of a polyubiquitin chain attached to the target protein, making it a preferred substrate for the 26S proteasome. The 26S proteasome, which hydrolyzes the targeted proteins, is composed of a 20S proteolytic core flanked by one or two 19S regulatory particles (17). The 19S regulatory particle functions to acquire ubiquitylated substrates and direct them into the proteolytic chamber (18, 40).

Proteasome inhibitors have been shown to possess strong anti-tumor activity and are used in the treatment of multiple myeloma. One such inhibitor, bortezomib (Velcade), was the first approved compound in this new category of cancer treatment (43). Proteasome inhibitor treatment can result in increased proteasome levels. Recently, an adaptive feedback mechanism was identified where long-term treatment of human lymphoma cells with bortezomib induced increased *de novo* biogenesis of proteasomes (14). This allowed the cells to survive proteasome inhibition and to become hyperproliferative and apotosis-resistant. A number of cancer cell types have been shown to have abnormally high proteasome levels, including certain human hematopoietic tumor cells (26). Identifying the factors that participate in transcription of proteasome subunit mRNAs would be valuable in understanding the regulation of ubiquitin proteasome activity and help explain the anti-tumor activity of proteasome inhibitors.

In *Saccharomyces cerevisiae*, a well-defined transcription negative-feedback loop controls proteasome levels (11, 31, 53). The transcription factor RPN4 binds to promoter PACE elements and promotes transcription of proteasome and related genes. RPN4 is a rapidly degraded proteasome

MCB Accepts published online ahead of print

substrate and is present at low levels when proteasome activity is sufficient for degradation of UPS substrates. Surprisingly, the RPN4 negative-feedback network has only been identified in *Saccharomyces cerevisiae* and the closely related yeast, *Hemiascomycetes* (30).

Treatment of mammalian or *Drosophila* cells with proteasome inhibitors results in the up-regulation of proteasome subunits (29, 34). In both *Drosophila* adults and cell culture the depletion of one of the proteasome ubiquitin receptors, S5a/Pros54 (PSMD4), strongly increases the specific transcription and over-production of proteasome subunits (29, 48). Importantly, *Drosophila* cells depleted of S5a do not up-regulate stress or heat shock genes, suggesting that a proteasomal gene-specific transcriptional regulatory pathway exists in *Drosophila*.

Genome wide RNAi libraries have proven to be a powerful tool to identify new essential genes in many pathways (9). We have used one such genome wide RNAi library of transcription related genes to identify the transcription factors that may control metazoan proteasome levels. Initial validations of our screen identify a transcription factor, Cnc-C, which contributes to the expression of proteasomal components in *Drosophila*. It has been shown previously that *Drosophila* Cnc-C plays a role in oxidative stress tolerance, similar to the mammalian Nrf2 transcription factor (nuclear factor-erythroid 2 p45 subunit-related factor 2), and it has been proposed to be a direct homolog of mammalian Nrf2 (46). In this work we show by phylogenetic analysis that duplication of the Nrf genes occurred during vertebrate evolution and we instead propose that Cnc-C resembles the common ancestral complex to the mammalian Nrf transcription factor gene family.

Heat stress results in strong up-regulation of proteasomal activity in human fibroblasts (2). However, it has been shown that proteasome genes are not co-regulated by the same transcription factor, Hsf1, as heat shock proteins in response to cellular stress in mammalian cells (49). Recently, the Nrf1 transcription factor (TCF11) was shown to be important for induction of proteasome gene transcription in mammalian cells (41, 45). Our current results demonstrate a functional similarity between the mammalian Nrf1

MCB Accepts published online ahead of print

transcription factor and *Drosophila* Cnc-C, suggesting that there is an evolutionary conserved function in proteasomal gene transcription for this specific transcription factor. Using a V5His tagged Cnc-C construct we show that the Cnc-C protein is degraded by the proteasome and is also stabilized by depletion of the proteasome ubiquitin receptor S5a. We present a model where both oxidative stress tolerance and proteasome induction are controlled by a single transcription factor in *Drosophila*, Cnc-C.

## Results

### Inhibition of $Ub^{G76V}$GFP proteasomal degradation screened through a genomic wide RNAi library of transcription factors

To screen for factors that are required for the specific transcription of proteasome genes, a *Drosophila* S2 stable cell line (29) expressing the ubiquitin fusion protein, $Ub^{G76V}$GFP (8), was used in an RNAi screen. The $Ub^{G76V}$GFP is a widely used proteasomal reporter where the single G76V mutation inhibits cleavage of the ubiquitin from the GFP and converts the GFP into a UFD proteasome substrate (Fig. 1A). $Ub^{G76V}$GFP is constitutively expressed at high levels via an actin promoter in these cells and results in low fluorescence levels when UPS activity is at wild-type levels. The $Ub^{G76V}$GFP reporter requires both the 26S proteasome and pre-processing by the p97 ATPase for degradation (3). We predicted that RNAi depletion of a factor that is specifically responsible for transcription of proteasome pathway genes, rather than a part of the general transcription machinery, would deplete the cell of proteasomes, but would not interfere with the ability to express the $Ub^{G76V}$GFP reporter.

The dsRNAs used in this screen were from a *Drosophila* transcription factor RNAi sub-library (DRSC TRXN) (42). This sub-library targets *Drosophila* genes that have either proven or predicted roles as transcription factors. The library contains on average two amplicons per gene and a total of 993 genes targeted. After 2 days of RNAi treatment the library was screened using inverted fluorescent microscopy and scored by two independent screeners for three consecutive days. Variable auto-fluorescence from the

MCB Accepts published online ahead of print

well surfaces compromised the results obtained by automatic fluorescent plate reading and therefore manual screening was used. Knockdown of the p97 ATPase or proteasome subunits were used to set a maximum positive score value. The assay used a blind screening method where the identity of the genes targeted by the dsRNA in the wells was unknown. Each day two screeners scored a 384 well plate. On following days the screeners were switched so that all examiners scored every RNAi well in the RNAi library at least once. The scores were averaged and compared in a heat map format (Fig. 1B and Sup. Fig. 1). For the heat map, only scores that were marked positive at some level by both independent screeners are shown. Initially, a total of 52 RNAi treatments resulted in increased levels of $Ub^{G76V}$GFP. The majority of the 52 targeted genes showed only minor stabilization of the UPS reporter while a small number showed stronger increases. The identified genes were grouped into three classes; group I, stabilization observed consistently over the three day screen period; group II, stabilization of $Ub^{G76V}$GFP reporter that occurred for two of the three days; group III, stabilization that occurred only on a single day. The class I group of RNAi targeted transcription factors had the highest levels of $Ub^{G76V}$GFP reporter stabilization.

For the majority of transcription factors screened, two distinct dsRNAs were present in the library. However, for the majority of candidates found to be positive only a single dsRNA led to visual stabilization of $Ub^{G76V}$GFP. The exceptions that showed multiple positive hits in the library screen were the *Drosophila* genes, *cnc*, *spen*, *cenp-C*, *foxo*, and the little studied genes CG34406, similar to the human zinc finger protein 84 (ZNF84) and CG5366, high similarity to the cullin-associated NEDD8-dissociated protein 1 TIP120/CAND. The proposed small co-subunit to Cnc, the Maf protein Maf-S, was also initially observed to stabilize $Ub^{G76V}$GFP. The small size of the *maf-S* gene meant that the RNAi library only contained a single amplicon for *maf-S*. The fact that the other candidates did not show consistent stabilization in different amplicon targets suggests that the majority of observed changes are likely the result of indirect effects , such as inhibition of general cellular

MCB Accepts published online ahead of print

homeostasis and blockage of turnover of the highly expressed reporter gene. A second possibility is that the difference seen between the amplicon pair for a single gene represents a real variation in the role of the alternative transcripts of that gene. Further characterization of the remaining candidate transcription factors is required to address false positives versus real contributors to maintaining proteasome levels.

To confirm that the candidate genes were required for the expression of genes of the ubiquitin proteasome system, larger scale RNAi targeting was carried out on the top positive candidate genes using dsRNA sequences that did not overlap the target sites from the DRSC TRXN sub-library. Depletion of *cnc*, *spen*, and *ssrp* were confirmed to stabilize the $Ub^{G76V}$GFP reporter. RNAi depletion of the Cnc transcription factor resulted in significant stabilization of the $Ub^{G76V}$GFP reporter (Fig. 1B and Sup. Fig. 2). Comparison of the specific regions targeted for the *cnc* gene in the original library screen indicated that only dsRNA directed to the *cnc*-C transcript was sufficient to inhibit proteasome degradation of $Ub^{G76V}$GFP, top hit in Fig. 1B and Sup. Fig. 2. Somewhat surprisingly, a second round of treatment with specific dsRNAs directed to *maf-S* or *kr-h1* was unable to show significant stabilization of the proteasome reporter. Previous studies suggest that Maf-S should be essential for Cnc-C to function as an obligate heterodimer. Additional studies are required to address the irreproducibility of the original hit for *maf-S* to understand if this is due to insufficient RNAi depletion or biological in nature.

**_Drosophila cnc_ gene resembles the ancestral gene for the mammalian Nrf transcription factors-diversified functions in a single locus**

The *Drosophila cnc* gene encodes seven different mRNAs (Fig. 2A). The majority of these transcriptional forms are due to different transcriptional start sites and not alternative splicing. The related mammalian homologs of *Drosophila cnc* include the p45 NFE2 (nuclear factor erythroid-derived 2) and the Nrf family of bZIP genes (NFE2-related factors, Nrf1, Nrf2, and Nrf3) (Fig. 2B) (47). These four related transcription factors typically function as transcriptional activators. Two additional related vertebrate bZIP factors,

MCB Accepts published online ahead of print

Bach1 and Bach2, function as transcriptional repressors (39). Similar to its mammalian homologs, the proteins encoded by the alternative transcripts of *cnc* have been proposed to serve various functions in *Drosophila*, with Cnc-A functioning as a putative repressor similar to the Bach factors; Cnc-B functioning as a developmental transcriptional regulator similar to p45 NFE2; and Cnc-C a response mediator similar to the Nrfs (47).

Our RNAi screen implicated the protein encoded by the longest transcript of *cnc*, *cnc-C*, in *Drosophila* proteasome regulation. RNAi depletion using dsRNA targeted specifically to *cnc-C* was equivalent or better at inhibiting proteasome degradation than dsRNA constructs that targeted all *cnc* transcripts. The *cnc-C* transcript contains a long unique 5' sequence that is absent in the other transcripts (Fig. 2A). Translation from this mRNA results in a Cnc-C protein containing 1383 amino acid residues, of which the N-terminal 578 are unique among the Cnc isoforms. The *Drosophila* Cnc-C protein has been proposed to have Nrf2-like functions and act as the central mediator of the antioxidant response system (46). This classification is based on the ability of Cnc-C to promote transcription of anti-oxidant genes, and the observation that the Cnc-C can also be regulated by a homologous cytoplasmic binding repressor, Keap1, as found with the vertebrate Nrf2 transcription factor. The vertebrate Nrf1 transcription factor also has Keap1 binding domains; however, membrane anchoring of Nrf1, rather than direct binding of Keap1, is likely to regulate Nrf1's ability to activate transcription.

In support of the idea that *Drosophila* Cnc has multiple transcriptional activities, comparison of conserved functional domains shows that Cnc-C contains both Nrf1 and Nrf2 domains (Fig. 2C). The Cnc-C protein shares sequence similarity with the Nrf1 N-terminus and also the Keap1 binding domains present in both Nrf2 and Nrf1. The function of Keap1 binding domains present in Nrf1 is currently unknown. It has been shown that the Nrf1 N-terminus is an important functional regulatory element that anchors the transcription factor to the membrane (57). The above analysis supports the idea that the various functions for the Nrf family were present before multiple duplications in modern vertebrates.

MCB Accepts published online ahead of print

While Cnc-C clearly has a role in antioxidant response (44), phylogenetic analysis does not support Cnc-C being a specific invertebrate homolog of the mammalian Nrf2 (Fig. 3). A comprehensive phylogenetic analysis with metazoan Cap-N-Collar bZIP sequences from a wide range of metazoan species supports an evolutionary model where the Nrf2 gene arose after the split between vertebrates and invertebrates. This is also supported by a previous study (25). Interestingly, the phylogenetic analysis also suggests that the duplication of bZIP proteins appears to have occurred in Chordata after the divergence of the vertebrates from the primitive chordates such as the Tunicates, *Ciona intestinalis,* an urocordate and the Lancelet, *Branchiostoma floridae*, a cephalocordate.

**RNAi depletion of *cnc-C* mRNA decreases levels of the 26S proteasome**

The stabilization of $Ub^{G76V}$GFP following gene-specific dsRNA knockdown can be caused by either a general decline in the expression of the whole 26S proteasome, of a proteasome sub-complex, or a single critical component. We attempted to identify the type of defects in the *Drosophila* UPS system resulting from the RNAi knockdowns of the top candidate transcription factors, focusing specifically on the roles of Cnc-C. Immunoblots of protein isolated from *Drosophila* cells harvested 3 days after treatment with dsRNA to *cnc-C* resulted in a decrease in individual proteasome subunits from different sub-complexes of the proteasome (Fig. 4A). The greatest decline at the protein level was observed for the 20S proteasome subunit alpha7 followed by the 19S subunit Rpt 1. Knockdown of *cnc-C* levels resulted in a modest decline of the p97 ATPase. As has been previously observed for *Drosophila* cells, RNAi knockdown of the ubiquitin receptor, S5a, resulted in increased levels of proteasome subunits. To examine the abundance of the proteasome complex itself, equivalent amounts of total cellular protein were separated by native gel electrophoresis, and in-gel peptidase assays were carried out by overlaying with the fluorogenic proteasome substrate LLVY-AMC, followed by exposure to UV light, (Fig. 4B). After knockdown of *cnc-C* levels the in-gel peptidase assays showed reductions of 26S proteasome complexes but little or no

MCB Accepts published online ahead of print

apparent decreases in the 20S proteasome. The cellular activity for 20S proteasomes was also examined with a more sensitive assay using whole cell protein extracts and following the cleavage of LLVY-AMC in a standard cuvette based assay (Fig. 4C). The 20S proteasome assays on RNAi treated cell extracts showed a strong reduction of the 20S proteasome activity when Cnc-C levels were depleted with different dsRNAs that were specific for the *cnc-C* transcripts. This assay measures total 20S proteasome pools both free 20S and 20S proteasome that were assembled in the 26S proteasome. The apparent discrepancy between the two 20S proteasome assays may be due to changes in the ratios of 20S versus 26S proteasomes. Inhibited assembly of 26S proteasomes would tend to maintain the levels of free 20S proteasomes as observed with the in-gel peptidase assay even though the overall level of 20S proteasomes has declined as measured with the cuvette based assay.

In addition, *cnc* RNAi could prevent the S5a subunit RNAi-related increase in 20S subunit activity. The S5a RNAi knockdown was effective in the presence of dsRNAs to *cnc*, and immunoblots against S5a showed apparently equivalent depletion in S5a when treated with both dsRNAs or alone, (Fig. 4A).

**Depletion of the Keap1 substrate adaptor protein does not induce proteasome level increases**

In *Drosophila* it has been shown that the expression of Keap1 rescues defects due to overexpression of *cnc*-C in the eye (46). In mammals Keap1 is a BTB-Kelch-type substrate adaptor protein of the Cul3-dependent ubiquitin ligase complex and functions as an adaptor for the Cullin 3 based E3 ligase to regulate the stability of the CNC protein Nrf2 (6, 15, 24, 55). As found in mammalian systems with Keap1 and Nrf2, *Drosophila* Keap1 and Cnc-C are proposed to directly bind to each other and have been identified to interact in two-hybrid experiments (16). In the standard model, Keap1 acts as a negative regulator of Cnc-C and together they play an important role in the cell as an important sensor for oxidants and toxic agents, during antioxidant and

MCB Accepts published online ahead of print

detoxification responses (see Fig. 7).

To examine the effects on proteasome subunit levels the *Drosophila* Keap1 was knocked down with RNAi (Fig. 4D). Surprisingly, no increase in any proteasome subunit levels was apparent after dsRNA treatment against Keap1, but showed a decrease in each subunit. Quantitative RT-PCR measurements of cells treated with dsRNA to *Keap1* showed that partial depletion of *Keap1* by RNAi treatment could significantly increase an oxidative stress response gene, *gstD1,* but no significant increase in mRNA level was observed for a proteasome subunit (Sup. Fig. 3).

**Cnc-C is essential for the *Drosophila* proteasome recovery pathway**

To confirm that RNAi directed to *cnc* led to decreases at the transcriptional level, mRNA levels for several components of the 26S proteasome were measured by quantitative RT-PCR (Fig. 5A). Levels of mRNA for one subunit each from the 20S proteasome (Alpha 7), the 19S regulatory base (Rpt3), and the 19S lid (Rpn11) were measured. The levels of the p97 ATPase mRNA were also measured, which have been shown to be co-regulated with the proteasome in *S. cerevisiae* and *Drosophila* (23, 29). The levels were normalized to the *RpL32* mRNA standard and compared to a control RNAi knockdown using dsRNA to GFP. Knockdown of these transcription factors resulted in modest but significant decreases at the mRNA level for all three classes of proteasome genes from 20-40% declines (Fig. 5A).

Consistent with the earlier RNAi screening, knocking down a single cnc transcript, *cnc-C*, resulted in decreases at the proteasome mRNA level that were equal or better than targeting all *cnc* transcripts. A significant decline in the level of mRNA for the p97 ATPase was also observed following *cnc-C* knock-down. A Taqman probe was available to examine the mRNA levels of *cnc* after RNAi treatments and confirmed specificity of the dsRNA treatments. The commercially available probe was designed to anneal to all the predicted transcript forms of cnc. Targeting specifically to the *cnc-C* transcript lowered the mRNA levels of the overall *cnc* mRNA less (cnc-C dsRNA#1 – 36%;

MCB Accepts published online ahead of print

dsRNA#2 – 29%) than targeting with one designed to target all *cnc* transcripts (cnc-All – 42%).

A number of ubiquitin binding proteins have been characterized and several others have been proposed to assist in the proteasomal degradation of ubiquitylated substrates. One of these facilitator proteins is the proteasome subunit S5a. Depletion of the ubiquitin binding subunit S5a increases the protein levels of *Drosophila* proteasome subunits (Fig. 4), but requires the presence of Cnc-C. Quantitative PCR measurements show that this regulation occurs at the mRNA level (Fig. 5B). As was seen at the protein level, the induction of proteasome mRNA after depletion of the ubiquitin receptor S5a was blocked in the absence of Cnc-C, indicating that the Cnc-C transcription factor is responsible for the increased levels of proteasome and p97 mRNAs after loss of the S5a proteasome subunit.

The recovery of proteasome levels after proteasome inhibitor treatment was also examined. After treatment with the proteasome inhibitor MG132 mammalian and *Drosophila* cells increase their proteasome component mRNA levels (29, 34). This has been termed the proteasome recovery pathway (41) and these increases in proteasome and p97 mRNA levels are blocked by RNAi against *cnc-C* (Fig. 5C). Overall these results show that Cnc-C is essential for the cell's ability to restore proteasome levels after proteasome inhibitor treatment and it contributes partially to maintaining basal levels of the proteasome mRNAs in steady state conditions.

**Cnc-C is a proteasome substrate and stabilized when the ubiquitin receptor S5a is depleted**

To understand the role of the proteasome in maintenance of Cnc-C protein levels a V5His-tagged recombinant Cnc-C containing 576 amino acid residues of the unique Cnc-C protein was expressed in *Drosophila* S2 cells. Since multiple attempts to generate polyclonal Cnc-C specific antibodies were unsuccessful, *cnc-C* was cloned from a *Drosophila* cDNA library and a C-terminal V5-His tag was added in place of the bZIP DNA binding domain whose sequence was not present in the EST transcript. A transfection of

MCB Accepts published online ahead of print

*Drosophila* S2 cells was divided and treated with either proteasome inhibitor or dsRNA and compared with untreated controls (Fig. 6A). Transfected control cells showed little or no detectable Cnc-C V5His protein; however, after epoxomicin treatment, the same transfection showed a high Cnc-C V5His level. Interestingly, RNAi knock-down of the S5a proteasome subunit, but not a GFP dsRNA treated control, also stabilized the Cnc-C V5His protein. As expected, RNAi treatment using the *cnc-C* dsRNA blocked the appearance of the Cnc-C V5His after proteasome inhibition.

It has been observed that the mammalian Nrf1 transcription factor is a membrane/ER bound protein and is proposed to be proteolytically cleaved for activation and transport to the nucleus independently of Keap1 (56). To determine if this is the case with *Drosophila* Cnc-C, *cnc-C* V5His transfected cells were harvested and membrane and soluble fractions were prepared and immunoblotted for the presence of the Cnc-C construct (Fig. 6B). A membrane bound fraction of the Cnc-C V5His had significantly slower migration rate through the SDS-PAGE gel than the soluble Cnc-C. Both the soluble and membrane Cnc-C V5His had significantly slower migration through the SDS-PAGE gel than expected by the size of the construct. Previously both human Nrf1 and Nrf2 transcription factors also have been observed to show abnormally slow migration in SDS-PAGE (56). The V5 antibody can only recognize aCnc-C protein that contains an intact C-terminus and this suggests that the small soluble form of Cnc-C present after proteasome inhibition may be caused by an N-terminal truncation. Post-translation modification such as glycosylation of the membrane form of Cnc-C cannot be currently ruled out to explain the migration differences. Assuming that the protease activity of the proteasome is inhibited after epoxomicin treatment, the proposed cleavage of the Cnc-C V5His would require a second protease for the generation of the soluble form. Additional studies are required to confirm and map the potential cleavage site(s) apparently present in the soluble Cnc-C V5His.

A number of bZIP transcription factors have been shown to be constitutive substrates for the 26S proteasome including Nrf2 (6). Our current results

MCB Accepts published online ahead of print

support a model where decreased proteasome activity or the loss of *Drosophila* S5a prevents the degradation of the Cnc-C transcription factor. Under conditions of low cellular proteasome activity the membrane bound Cnc-C protein is cleaved by a currently unknown proteolytic system and the liberation of a N-terminal truncated soluble form of Cnc-C promotes increased transcription of proteasome genes in the nucleus (Fig. 7).

## Discussion

Targeted inhibition of the 20S proteasome proteolytic activity has become a new and unique avenue in treating multiple myeloma and other cancers (44). While the understanding of how to inhibit the *in vivo* levels of the proteasome has greatly advanced in the past decade, less is known about the basic cellular mechanisms that increase or decrease overall proteasome levels, such as up-regulation of proteasome activity observed in muscle wasting conditions or the poorly understood age-related decline in proteasomes (7). To identify disease states that would be likely candidates for the use of proteasome inhibitors, it is important to understand the mechanisms and unique features of maintaining adequate proteasome levels in metazoan cells.

A widely used model substrate for measuring the UPS within living cells is the ubiquitin fusion, UFD substrate, $Ub^{G76V}$GFP (8) and was chosen as the reporter in this current RNAi screen. An advantage of using the $Ub^{G76V}$GFP reporter is that the cell assay requires the basic transcriptional machinery to still be functional, i.e. transcription off the actin promoter of the $Ub^{G76V}$GFP reporter has successfully occurred. Specific loss of transcription of UPS mRNAs is measured by the stabilization of the $Ub^{G76V}$GFP protein but does not create false positives for general transcription defects that would typically also prevent $Ub^{G76V}$GFP mRNA transcription.

Our current results indicate a role for the Cnc-C in maintaining 26S proteasome levels and the p97 ATPase. The *Drosophila* cnc gene was the first identified member of the extensive cap 'n' collar (CNC) family of basic leucine zipper proteins (36). The Cnc-bZIP factors are believed to function as obligate heterodimers with small Maf proteins (21). For *Drosophila* there is a

MCB Accepts published online ahead of print

single *cnc* gene that has a number of distinct transcript forms (33). The *cnc-C* transcript is expressed ubiquitously and is essential for life, while the *cnc-B* transcript is expressed in an embryonic pattern that includes the labral, intercalary and mandibular segments and is required for the proper development of these structures (33, 35, 51). A much more complex gene arrangement is present in mammalian cells represented by a family of bZIP factors related to the *Drosophila* Cnc protein and include the Nrf1, Nrf2, Nrf3, p45NFE2, and Bach1 and Bach2 transcription factors (47).

Both at the protein and mRNA level proteasome components in this study were found to decrease when *cnc-C* transcripts were depleted in *Drosophila* S2 cells. Overall, the decreases observed for the proteasome subunits mRNA levels were modest, 20-40% decline, when Cnc-C was depleted. The partial depletion of proteasome mRNA may indicate that transcription of proteasome mRNAs are regulated by multiple pathways and Cnc-C may only contribute a fractional role in the basal transcription of proteasome genes. Other candidates identified from our screen may also contribute to basal maintenance of proteasome mRNA levels. Surprisingly, large declines in the level of proteasome subunits were present in Cnc-C depleted cells that showed only a modest decline at the mRNA level (Figure 4C compared to Figure 5A). Over an extended period of time continuous sub-optimal mRNA levels may drain proteasome levels if proteasome turnover is significantly faster than replenishment.

On the other hand, the Cnc-C transcription factor does appear to have a dominant role in the re-establishment of proteasome levels after proteasome inhibitor treatment. A significant loss of proteasome mRNA induction was seen in cells depleted of Cnc-C. Also, the increase of proteasome mRNA levels caused by the depletion of the ubiquitin binding S5a protein was blocked when Cnc-C was depleted.

The *cnc* gene with its alternative transcripts may serve multiple roles in *Drosophila* similar to the multi-gene *nrf* transcription factor family in mammals. Treatment with dithiolethiones leads to increases of proteasome levels in mammalian cells through the oxidative stress pathway Nrf2-Keap1 and is due

MCB Accepts published online ahead of print

to increased transcription of genes for proteasome 20S and 19S subunits (27). However, it is unlikely that the Nrf2 plays a major role for the basal maintenance of 26S proteasome levels. Knockout of the nrf2 gene results in mice that are developmentally normal, but are more sensitive to chemical or environmental stress (5). In contrast with the knockout of *nrf2*, the knockout of the *nrf1* gene is embryonic lethal (4, 12). Nrf1 is essential for normal liver function in mice and liver-specific disruption of *nrf1* in mice results in increased inflammation and apoptosis and with additional time these mice spontaneously develop cancer (54).

*Drosophila* Cnc-C and mammalian Nrf1 N-terminal regions share sequence similarities and suggests possible similar regulation of the transcription factors. Nrf1 is believed not to be regulated by Keap1, but instead is proteolytically cleaved, liberating it from its membrane anchor, in a poorly understood process (57). Through the expression of a tagged form of a V5His Cnc-C, we also observe a possible proteolytic cleavage of a soluble form of the *Drosophila* transcription factor. Based on the retention of the C-terminal tag present in the expressed transgenic Cnc-C sequence, the cleavage likely occurs N-terminally. During the revision of our manuscript, the long isoform of human Nrf1, TCF11, was shown to be an essential regulator for 26S proteasome formation via an ERAD membrane feedback loop (45). For human TCF11 to upregulate proteasome subunit genes, both deglycosylation and cleavage of TCF11 is required for translocation from the ER membrane to the nucleus.

Cnc-C may control proteasome levels independently of Keap1. Previous *Drosophila* genomic transcriptional profiling experiments show that S5a depletion results in a specificity for inducing proteasome genes (29). Secondly, Keap1, which binds to and prevents Cnc-C from entering the nucleus and up-regulating anti-oxidant genes, is surprisingly one of the few non-proteasome genes transcriptionally up-regulated after S5a depletion when transcription of proteasome components is induced. Finally, RNAi depletion of *Keap1* does not increase proteasome levels as would be expected if Keap1 functioned to down-regulate proteasome transcription in

MCB Accepts published online ahead of print

*Drosophila* cells.

If *Drosophila* proteasome genes are under the control of the transcription factor Cnc-C it should be expected that proteasome genes share related promoter elements as found in other known genes that are transcriptionally activated by Cnc-C. Antioxidant response elements (ARE) are enhancer sequences that allow Nrf transcription factors to bind at the promoter region of regulated genes (22, 50). Recently, an ARE sequence has been identified as a regulatory element near the start of transcription for the mammalian proteasome gene PSMB6 (41). For *Drosophila* Cnc-C an ARE enhancer element has been identified for oxidative stress regulation and Cnc-C binding upstream of the glutathione S transferase gene, gstD (46). This *Drosophila* gstD regulatory sequence (TCAgcATGACcggGCAaaaa) shows clear similarity with the extended consensus of the mammalian ARE sequence motif (37). We have previously shown that increased levels of proteasome gene expression are regulated by small regions that overlap the transcriptional start regions of proteasome genes. For the proteasome genes Rpn1 and Beta-2, the transcriptional start regions contain ARE-like sequences, (GCAgtGTGACcgcGCGgcga) and (GAGcgATGACaaaCAAaatt), centered within previously mapped locations essential for induced transcription after proteasome inhibition (29). However, these sequences are directly after the transcription start site and are present on the anti-sense strand of the DNA (Sup. Fig. 4). Future studies will be required to understand the function of Cnc-C binding both oxidative stress regulated genes and proteasome genes.

Our recent results have led us to propose a speculative model for *Drosophila* proteasome regulation (Fig. 7). The model proposes that RNAi knockdown of the ubiquitin receptor protein S5a stabilizes Cnc-C, which is post-translationally modified and cleaved, to specifically serve in transcribing proteasome and proteasome related genes. It has been previously shown that loss of S5a stabilizes a subset of important cellular proteasome substrates. In *Saccharomyces cerevisiae* deletion of S5a (Rpn10) stabilizes a range of ubiquitin proteasome substrates including the bZIP transcription factor Gcn4 (32, 52).

MCB Accepts published online ahead of print

We propose that for *Drosophila,* a single protein, Cnc-C, is important for transcription of two large pathways of genes, anti-oxidant genes and proteasome genes and supports the idea that transcriptional specificity is occurring even though a single transcription factor is involved. Our model does not rule out the possibility that both pathways can be co-transcriptionally upregulated by Cnc-C. In fact, previous transcriptional array profiling of *Drosophila* cells recovering from proteasome inhibitor treatment shows upregulation of a number of anti-oxidant gene and xenobiotic-metabolizing gene mRNAs (29), including a 62-fold increase of a cytochrome P450, CYP6A20 and a 2-fold increase for the known Cnc-C regulated (46) glutathione S transferase D1. Additional work is required to understand the role of Cnc-C in *Drosophila* cells and the possible mechanisms to allow specific induction of the proteasome pathway. In conclusion, *Drosophila* Cnc appears to share an evolutionary conserved proteasome regulation pathway with the human Nrf transcription factors. Future examination of this regulatory-transcription pathway should help to explain how metazoan cells maintain proteasome activity and also identify additional targets for future therapeutic inhibitors of the proteasome system.

## References


1. Anisimova, M., and O. Gascuel. 2006. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Syst Biol 55:539-52.
2. Beedholm, R., B. F. Clark, and S. I. Rattan. 2004. Mild heat stress stimulates 20S proteasome and its 11S activator in human fibroblasts undergoing aging in vitro. Cell Stress Chaperones 9:49-57.
3. Beskow, A., K. B. Grimberg, L. C. Bott, F. A. Salomons, N. P. Dantuma, and P. Young. 2009. A conserved unfoldase activity for the p97 AAA-ATPase in proteasomal degradation. J Mol Biol 394:732-46.
4. Chan, J. Y., M. Kwong, R. Lu, J. Chang, B. Wang, T. S. Yen, and Y. W. Kan. 1998. Targeted disruption of the ubiquitous CNC-bZIP transcription factor, Nrf-1, results in anemia and embryonic lethality in mice. EMBO J 17:1779-87.
5. Chan, K., R. Lu, J. C. Chang, and Y. W. Kan. 1996. NRF2, a member of the NFE2 family of transcription factors, is not essential for murine erythropoiesis, growth, and development. Proc Natl Acad Sci U S A 93:13943-8.
6. Cullinan, S. B., J. D. Gordan, J. Jin, J. W. Harper, and J. A. Diehl. 2004. The Keap1-BTB protein is an adaptor that bridges Nrf2 to a Cul3-based E3 ligase: oxidative stress sensing by a Cul3-Keap1 ligase. Mol Cell Biol 24:8477-86.
7. Dahlmann, B. 2007. Role of proteasomes in disease. BMC Biochem 8 Suppl 1:S3.
8. Dantuma, N. P., K. Lindsten, R. Glas, M. Jellne, and M. G. Masucci. 2000. Short-lived green fluorescent proteins for quantifying ubiquitin/proteasome-dependent proteolysis in living cells. Nat Biotechnol 18:538-43.

MCB Accepts published online ahead of print

9. DasGupta, R., and F. C. Gonsalves. 2008. High-throughput RNAi screen in Drosophila. Methods Mol Biol 469:163-84.
10. Do, C. B., M. S. Mahabhashyam, M. Brudno, and S. Batzoglou. 2005. ProbCons: Probabilistic consistency-based multiple sequence alignment. Genome Res 15:330-40.
11. Dohmen, R. J., I. Willers, and A. J. Marques. 2007. Biting the hand that feeds: Rpn4-dependent feedback regulation of proteasome function. Biochim Biophys Acta 1773:1599-604.
12. Farmer, S. C., C. W. Sun, G. E. Winnier, B. L. Hogan, and T. M. Townes. 1997. The bZIP transcription factor LCR-F1 is essential for mesoderm formation in mouse development. Genes Dev 11:786-98.
13. Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27:401-410.
14. Fuchs, D., C. Berges, G. Opelz, V. Daniel, and C. Naujokat. 2008. Increased expression and altered subunit composition of proteasomes induced by continuous proteasome inhibition establish apoptosis resistance and hyperproliferation of Burkitt lymphoma cells. J Cell Biochem 103:270-83.
15. Furukawa, M., and Y. Xiong. 2005. BTB protein Keap1 targets antioxidant transcription factor Nrf2 for ubiquitination by the Cullin 3-Roc1 ligase. Mol Cell Biol 25:162-71.
16. Giot, L., J. S. Bader, C. Brouwer, A. Chaudhuri, B. Kuang, Y. Li, Y. L. Hao, C. E. Ooi, B. Godwin, E. Vitols, G. Vijayadamodar, P. Pochart, H. Machineni, M. Welsh, Y. Kong, B. Zerhusen, R. Malcolm, Z. Varrone, A. Collis, M. Minto, S. Burgess, L. McDaniel, E. Stimpson, F. Spriggs, J. Williams, K. Neurath, N. Ioime, M. Agee, E. Voss, K. Furtak, R. Renzulli, N. Aanensen, S. Carrolla, E. Bickelhaupt, Y. Lazovatsky, A. DaSilva, J. Zhong, C. A. Stanyon, R. L. Finley, Jr., K. P. White, M. Braverman, T. Jarvie, S. Gold, M. Leach, J. Knight, R. A. Shimkets, M. P. McKenna, J. Chant, and J. M. Rothberg. 2003. A protein interaction map of Drosophila melanogaster. Science 302:1727-36.
17. Glickman, M. H., and A. Ciechanover. 2002. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. Physiol Rev 82:373-428.
18. Groll, M., M. Bajorek, A. Kohler, L. Moroder, D. M. Rubin, R. Huber, M. H. Glickman, and D. Finley. 2000. A gated channel into the proteasome core particle. Nat Struct Biol 7:1062-7.
19. Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696-704.
20. Hordijk, W., and O. Gascuel. 2005. Improving the efficiency of SPR moves in phylogenetic tree search methods based on maximum likelihood. Bioinformatics 21:4338-47.
21. Itoh, K., K. Igarashi, N. Hayashi, M. Nishizawa, and M. Yamamoto. 1995. Cloning and characterization of a novel erythroid cell-derived CNC family transcription factor heterodimerizing with the small Maf family proteins. Mol Cell Biol 15:4184-93.
22. Johnsen, O., P. Murphy, H. Prydz, and A. B. Kolsto. 1998. Interaction of the CNC-bZIP factor TCF11/LCR-F1/Nrf1 with MafG: binding-site selection and regulation of transcription. Nucleic Acids Res 26:512-20.
23. Kapranov, A. B., M. V. Kuriatova, O. V. Preobrazhenskaia, V. V. Tiutiaeva, R. Shtuka, H. Feldmann, and V. L. Karpov. 2001. [Isolation and identification of PACE-binding protein rpn4--a new transcription activator, participating in regulation of 26S proteosome and other genes]. Mol Biol (Mosk) 35:420-31.
24. Kobayashi, A., M. I. Kang, H. Okawa, M. Ohtsuji, Y. Zenke, T. Chiba, K. Igarashi, and M. Yamamoto. 2004. Oxidative stress sensor Keap1 functions as an adaptor for Cul3-based E3 ligase to regulate proteasomal degradation of Nrf2. Mol Cell Biol 24:7130-9.
25. Kobayashi, M., K. Itoh, T. Suzuki, H. Osanai, K. Nishikawa, Y. Katoh, Y. Takagi, and M. Yamamoto. 2002. Identification of the interactive interface and phylogenic conservation of the Nrf2-Keap1 system. Genes Cells 7:807-20.
26. Kumatori, A., K. Tanaka, N. Inamura, S. Sone, T. Ogura, T. Matsumoto, T. Tachikawa, S. Shin, and A. Ichihara. 1990. Abnormally high expression of proteasomes in human leukemic cells. Proc Natl Acad Sci U S A 87:7071-5.
27. Kwak, M. K., N. Wakabayashi, J. L. Greenlaw, M. Yamamoto, and T. W. Kensler. 2003. Antioxidants enhance mammalian proteasome expression through the Keap1-Nrf2 signaling pathway. Mol Cell Biol 23:8786-94.

MCB Accepts published online ahead of print

28. Le, S. Q., and O. Gascuel. 2008. An improved general amino acid replacement matrix. Mol Biol Evol 25:1307-20.
29. Lundgren, J., P. Masson, Z. Mirzaei, and P. Young. 2005. Identification and characterization of a Drosophila proteasome regulatory network. Mol Cell Biol 25:4662-75.
30. Mannhaupt, G., and H. Feldmann. 2007. Genomic evolution of the proteasome system among hemiascomycetous yeasts. J Mol Evol 65:529-40.
31. Mannhaupt, G., R. Schnall, V. Karpov, I. Vetter, and H. Feldmann. 1999. Rpn4p acts as a transcription factor by binding to PACE, a nonamer box found upstream of 26S proteasomal and other genes in yeast. FEBS Lett 450:27-34.
32. Mayor, T., J. Graumann, J. Bryan, M. J. MacCoss, and R. J. Deshaies. 2007. Quantitative profiling of ubiquitylated proteins reveals proteasome substrates and the substrate repertoire influenced by the Rpn10 receptor pathway. Mol Cell Proteomics 6:1885-95.
33. McGinnis, N., E. Ragnhildstveit, A. Veraksa, and W. McGinnis. 1998. A cap 'n' collar protein isoform contains a selective Hox repressor function. Development 125:4553-64.
34. Meiners, S., D. Heyken, A. Weller, A. Ludwig, K. Stangl, P. M. Kloetzel, and E. Kruger. 2003. Inhibition of proteasome activity induces concerted expression of proteasome genes and de novo formation of Mammalian proteasomes. J Biol Chem 278:21517-25.
35. Mohler, J., J. W. Mahaffey, E. Deutsch, and K. Vani. 1995. Control of Drosophila head segment identity by the bZIP homeotic gene cnc. Development 121:237-47.
36. Mohler, J., K. Vani, S. Leung, and A. Epstein. 1991. Segmentally restricted, cephalic expression of a leucine zipper gene during Drosophila embryogenesis. Mech Dev 34:3-9.
37. Nioi, P., M. McMahon, K. Itoh, M. Yamamoto, and J. D. Hayes. 2003. Identification of a novel Nrf2-regulated antioxidant response element (ARE) in the mouse NAD(P)H:quinone oxidoreductase 1 gene: reassessment of the ARE consensus sequence. Biochem J 374:337-48.
38. Nouhi, Z., G. Chevillard, A. Derjuga, and V. Blank. 2007. Endoplasmic reticulum association and N-linked glycosylation of the human Nrf3 transcription factor. FEBS Lett 581:5401-6.
39. Oyake, T., K. Itoh, H. Motohashi, N. Hayashi, H. Hoshino, M. Nishizawa, M. Yamamoto, and K. Igarashi. 1996. Bach proteins belong to a novel family of BTB-basic leucine zipper transcription factors that interact with MafK and regulate transcription through the NF-E2 site. Mol Cell Biol 16:6083-95.
40. Rabl, J., D. M. Smith, Y. Yu, S. C. Chang, A. L. Goldberg, and Y. Cheng. 2008. Mechanism of gate opening in the 20S proteasome by the proteasomal ATPases. Mol Cell 30:360-8.
41. Radhakrishnan, S. K., C. S. Lee, P. Young, A. Beskow, J. Y. Chan, and R. J. Deshaies. 2010. Transcription factor Nrf1 mediates the proteasome recovery pathway after proteasome inhibition in mammalian cells. Mol Cell 38:17-28.
42. Ramadan, N., I. Flockhart, M. Booker, N. Perrimon, and B. Mathey-Prevot. 2007. Design and implementation of high-throughput RNAi screens in cultured Drosophila cells. Nat Protoc 2:2245-64.
43. Richardson, P. G., T. Hideshima, and K. C. Anderson. 2003. Bortezomib (PS-341): a novel, first-in-class proteasome inhibitor for the treatment of multiple myeloma and other cancers. Cancer Control 10:361-9.
44. Shah, J. J., and R. Z. Orlowski. 2009. Proteasome inhibitors in the treatment of multiple myeloma. Leukemia.
45. Steffen, J., M. Seeger, A. Koch, and E. Kruger. 2010. Proteasomal degradation is transcriptionally controlled by TCF11 via an ERAD-dependent feedback loop. Mol Cell 40:147-58.
46. Sykiotis, G. P., and D. Bohmann. 2008. Keap1/Nrf2 signaling regulates oxidative stress tolerance and lifespan in Drosophila. Dev Cell 14:76-85.
47. Sykiotis, G. P., and D. Bohmann. 2010. Stress-activated cap'n'collar transcription factors in aging and human disease. Sci Signal 3:re3.
48. Szlanka, T., L. Haracska, I. Kiss, P. Deak, E. Kurucz, I. Ando, E. Viragh, and A. Udvardy. 2003. Deletion of proteasomal subunit S5a/Rpn10/p54 causes lethality, multiple mitotic

MCB Accepts published online ahead of print

defects and overexpression of proteasomal genes in Drosophila melanogaster. J Cell Sci 116:1023-33.
49. Taylor, D. M., E. Kabashi, J. N. Agar, S. Minotti, and H. D. Durham. 2005. Proteasome activity or expression is not altered by activation of the heat shock transcription factor Hsf1 in cultured fibroblasts or myoblasts. Cell Stress Chaperones 10:230-41.
50. Venugopal, R., and A. K. Jaiswal. 1996. Nrf1 and Nrf2 positively and c-Fos and Fra1 negatively regulate the human antioxidant response element-mediated expression of NAD(P)H:quinone oxidoreductase1 gene. Proc Natl Acad Sci U S A 93:14960-5.
51. Veraksa, A., N. McGinnis, X. Li, J. Mohler, and W. McGinnis. 2000. Cap 'n' collar B cooperates with a small Maf subunit to specify pharyngeal development and suppress deformed homeotic function in the Drosophila head. Development 127:4023-37.
52. Verma, R., R. Oania, J. Graumann, and R. J. Deshaies. 2004. Multiubiquitin chain receptors define a layer of substrate selectivity in the ubiquitin-proteasome system. Cell 118:99-110.
53. Xie, Y., and A. Varshavsky. 2001. RPN4 is a ligand, substrate, and transcriptional regulator of the 26S proteasome: a negative feedback circuit. Proc Natl Acad Sci U S A 98:3056-61.
54. Xu, Z., L. Chen, L. Leung, T. S. Yen, C. Lee, and J. Y. Chan. 2005. Liver-specific inactivation of the Nrf1 gene in adult mouse leads to nonalcoholic steatohepatitis and hepatic neoplasia. Proc Natl Acad Sci U S A 102:4120-5.
55. Zhang, D. D., S. C. Lo, J. V. Cross, D. J. Templeton, and M. Hannink. 2004. Keap1 is a redox-regulated substrate adaptor protein for a Cul3-dependent ubiquitin ligase complex. Mol Cell Biol 24:10941-53.
56. Zhang, Y., D. H. Crouch, M. Yamamoto, and J. D. Hayes. 2006. Negative regulation of the Nrf1 transcription factor by its N-terminal domain is independent of Keap1: Nrf1, but not Nrf2, is targeted to the endoplasmic reticulum. Biochem J 399:373-85.
57. Zhang, Y., J. M. Lucocq, M. Yamamoto, and J. D. Hayes. 2007. The NHB1 (N-terminal homology box 1) sequence in transcription factor Nrf1 is required to anchor it to the endoplasmic reticulum and also to enable its asparagine-glycosylation. Biochem J 408:161-72.

## Materials and Methods

### RNAi interference library screen—Screening of 384-well plates

384-well plates with pre-aliquoted double-stranded RNA (dsRNA) were thawed for a few minutes and then spun down at 200 x g for 1 minute. $Ub^{G76V}$GFP stable cells were counted, spun down at 100 x g for 5 minutes and washed twice, after which they were resuspended in serum-free media at a concentration of $2.5x10^6$ cell/mL. The adhesive seal on the plates was removed and 10 $\mu$L cells were plated in each well. The plates were incubated for 2 hours, after which 30 $\mu$L complete media were added to each well. The plates were sealed and incubated for 4 days, with screening on day 2, 3 and 4 for each plate. Each plate was screened each day by two persons, each well given a score of 0 to 4, with 0 being no stabilization of $Ub^{G76V}$GFP and 4 being high stabilization of $Ub^{G76V}$GFP.

MCB Accepts published online ahead of print

**Follow-up screen—**

Using the MEGAscript kit from Ambion, dsRNA for positives knockdowns from the 1st and 2nd screening was synthesized. Oligonucleotides with T7 RNA polymerase promoter regions were constructed from gene sequences found in FlyBase. Fragments of approximately 500-700 bp were amplified using *Drosophila* genomic DNA extracted from S2 cells, after which dsRNA was synthesized from fragments. All dsRNA were examined by agarose gel electrophoresis after annealing. Oligonucleotide sequences used to generate fragments were as follows:

T7 5' cncC#1 5'-GAATTAATACGACTCACTATAGGGAGAGGCTGCAAGCTTCCGCCAAGATTCAACG
T7 3' cncC#1 5'-GAATTAATACGACTCACTATAGGGAGAGTGCTGAGGGGTGCTCCACTGCCGC
T7 5' cncC#2 5'-GAATTAATACGACTCACTATAGGGAGAATTGGAAAATCGGGGAGATAGGCCGTGG
T7 3' cncC#2 5'-GAATTAATACGACTCACTATAGGGAGAACAAAAGTGCTGCCGTTGAATCTTGGCG
T7 5' cnc-All 5'-GAATTAATACGACTCACTATAGGGAGAGGCTACAGCTGCCTCCAATGCTGTTTCG
T7 3' cnc-All 5'-GAATTAATACGACTCACTATAGGGAGATCCAGAAGGGACAACACCATGCAAGACG
T7 5' ssrp:5'-GAATTAATACGACTCACTATAGGGAGAGCCCGCAGCGGCGGATCC-3'
T7 3' ssrp:5'-GAATTAATACGACTCACTATAGGGAGAGCTCCTTCCACATCTCGCCGCCC-3'
T7 5' spen: 5'-GAATTAATACGACTCACTATAGGGAGAGGGCTTTGGAAAATCCATGCCCACC-3'
T7 3' spen:5'-GAATTAATACGACTCACTATAGGGAGAGGGCATGGATGTGCTCGCTGGTACC-3'
T7 5' cenp-c:5'-GAATTAATACGACTCACTATAGGGAGAGCCGCCACCGAGAAAGTCAATGAGC-3'
T7 3' cenp-c:5'-GAATTAATACGACTCACTATAGGGAGAGCCAGGCTGGGTCGTGACTTTGTCG-3'
T7 5' kr-h1:5'-GAATTAATACGACTCACTATAGGGAGAGGTGTCGCAGGTGCGCAAGCCC-3'
T7 3' kr-h1:5'-GAATTAATACGACTCACTATAGGGAGACGGGCACTTGTACGGCTTCTCGCC-3'

For knockdown, *Drosophila* S2 cells were diluted to a concentration of $1x10^6$ cells/ml in Express Five Serum Free media (Invitrogen). After cell attachment, medium was replaced twice with 1 ml serum-free media, and the cells were grown in the serum-free media for 1 to 2 h. dsRNA were introduced into cells by adding 20-60 $\mu$g of dsRNA directly to the medium under constant agitation. After 12 h, 2 ml complete S2 media (Invitrogen) were added, and the cells were cultured for 3 to 4 days prior to isolation.

**Synthesis of dsRNA—**

A second round of double stranded RNAs (approximately 600 bp length) were generated to candidate transcription factors and to proteasome genes. For each treatment, a total of 5 $\mu$g of dsRNA was supplemented to the media of a

MCB Accepts published online ahead of print

stable Drosophila cell line expressing the UFD reporter UbG76V-GFP using 96 well format. Triplicates were carried out for controls and all RNAi treatments. After three days of growth the cells were measured by flow cytometry, FACs.

**Fluorometric Assays of Proteasome Activities**—Spectrofluorometric assays were performed in the presence of Suc-LLVY-AMC, succinyl-Leu-Leu-Val-Tyr-AMC succinyl-Leu-Leu-Val-Tyr-7-amino-4-methyl-coumarin, and 10 μg of each crude extract from RNAi treated Drosophila S2 extracts in 100 μl TS buffer (10 mM Tris-HCl. pH 8.5, 25 mM KCl, 10 mM NaCl, 1.1 mM MgCl2, and 0.1 mM EDTA). Before initiating the proteasome assays equivalent amounts of extracts were diluted into TS buffer with 0.03 % SDS present and pre-incubated 15 minutes at room temperature in the absence of ATP to allow activation of the 20S proteasomes. Concentration of fluorescent peptide was 200 μM. Assay volumes were 100 μl and incubated at 37 C for 1 hr. Assays were stopped by the addition of 1 ml ice cold water. Fluorescence was measured using an excitation wavelength of 380 nm and an emission wavelength of 440 nm. Suc-LLVY-AMC was purchased from AFFINITI Research Products.

**RNAi Knockdowns and RNA purifications**—
Day one, Schneider's *Drosophila* medium cultured S2 cells were plated on Corning costar 6 well plates ~ 2,25*$10^6$ cells in each well. The cells were washed 4 times with Express Five serum-free medium (Invitrogen) with 1-hour incubation in between washes. A total of 20 μg dsRNA were applied to the cells now sitting in 1 mL serum-free media. When two different dsRNA were added, 2*20 μg dsRNA were used. Day two, 2 mL Schneider's media was added to each well. Day three, total RNA was isolated with TRIzol® Plus RNA Purification Kit (Invitrogen).

MCB Accepts published online ahead of print

**Reverse Transcription and Quantitative PCR—**
First, the isolated RNA was treated with DNAse (Turbo DNA-free kit from Ambion). To make cDNA, the high capacity cDNA Reverse Transcription Kit was used (Applied Biosystems) and 490 ng RNA were used in a 20 μL reaction. The resulting cDNA was then used in QPCR reactions that was performed with TaqMan gene expression Assay. The following Taqman pre-designed probes were used: Drosophila prosalpha7 (Dm01812483_g1), Rpt3 (Dm01837846_g1), cnc (Dm02150448_g1), p97 (TER94) (Dm01812530_g1), and Rpn11(Dm01804595_g1). Ribosomal protein L32 (Dm01812483_g1) was used as an internal control for all measurements. Each of the three biological replicated experiments was examined in triplicates. Optimized amounts of cDNA were originally tested for the samples to measure the different gene levels. One individual sample of the triplicate for maf-s and cnc contained 85 ng cDNA, for prosalpha 7, Rpt3, p97 and Rpn11 27 ng was used and for the ribosomal protein, 54 ng were used. The QPCR reactions were performed on the ABI Prism 7000 (Applied Biosystems) apparatus and the data were analyzed according to the ΔCt method.

**Phylogenic tree of cap-n-collar proteins–** Sequences were collected from GenBank with PSI -BLAST using the *Drosophila melanogaster* Cnc isoform A protein sequence (NCBI accession number NP_732835) as seed and an E-value cutoff of 1e-40. The dataset was reduced by deletion of sequences from closely related species and isoforms that were very similar to other sequences in the dataset. Subsequently, the dataset was complemented with sequences from *Caenorhabditis* spp., *Nematostella cectensis* and *Trichoplax adhaerens* that did not pass the E-value cutoff of the initial run. A multiple alignment was constructed using Probcons 1.10 (10) and phylogentic trees were estimated using PhyML 3.0 (19, 20).

**Expressing a V5/His tagged form of the unique Cnc-C region–** The cDNA iPCR clone IP15234 was obtained from the *Drosophila* Genomics Resource Center and fully sequenced to confirm the open reading frame. The partial

MCB Accepts published online ahead of print

cDNA contained the complete region for the Cnc-C specific region of Cnc but lacked the bZIP DNA binding domain. Through PCR an open reading frame sequence was obtained with flanking EcoR1 and NotI restriction sites and cloned into the EcoR1/NotI sites of pAcV5His-B (Invitrogen) plasmid that trnalsates into a C-terminally V5-tagged protein. The oligos used for the PCR were the following:

5'-ATTAAGAATTCGACTACAAGAGTCACCCACGCACCCATTCGC-3'

5'-TAATTAGCGGCCGCCCTCGTCTAGTTCCTTCTTGGCGTCCTCATCTAAG-3'

The resulting clone was sequenced and site directed mutagenesis was carried out to fix two unwanted mutations present in the PCR clone.

Oligos used for mutagensis were:

V309I 5'-GATTACGAGGGCGAGCTGATCGGTGGAGTGGCCAACG-3'

K517E 5'-CGGATTGTTTCGGAAACCGGCGAGGATTTACTCAGTGGC-3'

The resulting clone expressed a 611 amino acid residue length C-terminal tag V5-His C region. The Cnc-C region itself was 576 amino acid residues in length and only contained Cnc-C specific sequence and not the bZIP binding domain found in all Cnc transcripts.

## Figures

Figure 1. A. Overview of the genome-wide RNAi screen to identify transcription factors involved in regulating the ubiquitin proteasome system. The *Drosophila* Transcription Factor RNAi Sub-Library (DRSC TRXN) was manually screened using a stable S2 cell line. A total of 993 gene target wells were examined for the library screen. A *Drosophila* S2 stable cell line that constitutively expressed $Ub^{G76V}$GFP reporter was used in the library screening and is not detectable in *Drosophila* cells that have adequate proteasome levels but is observed in cells with depleted proteasomes levels. B. Ranking for stabilizing the proteasome reporter $Ub^{G76V}$GFP from a RNAi genomic wide screening of *Drosophila* transcription factors. The positive hits were ranked in a heat map and grouped into categories based on the number of days that $Ub^{G76V}$GFP was observed to be stabilized. Group I represents targets that showed consistent stabilization for all three screening periods. Group II genes stabilized $Ub^{G76V}$GFP for two consecutive days. Brackets show target genes

MCB Accepts published online ahead of print

that were found twice in the library screen with different target dsRNAs. The strongest stabilization of $Ub^{G76V}$GFP was found with a cnc dsRNA target that mapped to a specific transcript of the gene, cnc-C, shown as the top row. The top candidate genes were re-targeted with new dsRNA treatments to new exonic regions. Triplicate treatments were analyzed by flow cytometry and measured for $Ub^{G76V}$GFP stabilization compared to controls. For the cnc gene retargeting specifically to unique sequences specific for cnc-C transcript was functional in stabilizing $Ub^{G76V}$GFP. Student's *t*-test value (P = 0.0004) is shown.

Figure 2. Comparison of the cnc family of transcription factors in mammals versus *Drosophila* A. Schematic diagram of alternative splice forms of the basic-leucine zipper gene *cnc*. Translated regions are shown as gray boxes while UTR regions are shown as white filled boxes. The positions of the dsRNA sequences from the initial library are shown as short lines and confirmative dsRNAs are shown as red lines. Double stranded RNA was designed to target all alternative splice forms, -All or specific for the Cnc-C transcripts. B. At the protein level the ETGE domain is part of a conserved region that binds to Keap1 homologs and prevents import to the nucleus for the Nrf2 in mammalian cells and is also proposed to function in binding *Drosophila* Cnc-C during down regulation of anti-oxidant gene transcription. The mammalian system contains three distinct genes Nrf1, Nrf2 and Nrf3. C. The N-termini of Nrf1 and Nrf3 function as signal peptide sequences (38, 57) and anchor the proteins to membranes. The *Drosophila* Cnc-C N-terminus shows similarities to both mammalian Nrf1 and Nrf2 critical domains. The *Drosophila* Cnc-C and other insect Cnc sequences show similarity to the mammalian Nrf1 N-terminal region, which has been shown to target Nrf1 to the ER and cell membranes but is absent in Nrf2 transcription factors. The *Drosophila* Cnc-C also contains two positionally conserved Keap1 binding domains that target mammalian Nrf2 to the 26S proteasome.

MCB Accepts published online ahead of print

Figure 3. Phylogeny of Cnc and Nrf proteins from metazoa rooted with the most basal lineage included, *Trichoplax adhaerens*. The tree was estimated using PhyML (19, 20) using the LG substitution model (28) and four-category gamma correction of different mutation rates along sequences. Support values shown are the SH-like values described (1). Except for the grouping of *Caendorhabditis* spp. with *Menatostella vectensis* (starlet sea anemone, a non-bilaterian) all metazoan groups show up in expected places of the tree. In particular, the two species most closely related to vertebrates, *Ciona intestinalis* (vase tunicate, a urocordate) and *Branchiostoma floridae* (lancelet, a cephalocordate) form a sister group with vertebrates. The latter contains four paralogs, while all other species contain just a single gene from the group. It can then be concluded that the ancestral gene containing the conserved domain has been duplicated more than once in the vertebrate lineage, and all four paralogs in vertebrates are equally closely related to for instance insect Cnc genes. The position of the *Caenorhabditis* spp. sequences can possibly be an artifact caused by sequence divergence (13). Furthemore, SKN sequences for *Caenorhabditis* contain a conserved domain that is shorter than all other sequences in the tree.

Figure 4. Decrease in the expression of proteasome subunits after knockdown of Cnc-C. A. Immunoblots to subunits from different sub-complexes of the 26S proteasome and p97 ATPase. RNAi knockdown of ubiquitin receptor S5a has been previously shown to increase overall 26S proteasome levels and the p97 ATPase (29). Co-RNAi depletions of S5a and Cnc blocks S5a induced increases in proteasome levels. B. Results from a native gel proteasome assay where 26S and 20S proteasomes are separated by electrophoresis and identified by in-gel hydrolysis of LLVY-AMC proteasome fluorogenic substrate. C. 20S proteasome assay of Cnc-C RNAi depleted Drosophila extracts. After RNAi treatments the cell extracts were measured for total protein concentrations and assayed for 20S proteasome levels. Assays were carried out in the absence of ATP and the addition of dilute SDS (0.03%) to activate the 20S proteasomes. Activity determined by hydrolysis of Suc-LLVY-AMC

MCB Accepts published online ahead of print

(final concentration 200 $\mu$M) for 1 hr at 37°C. D. Immunoblots to 26S proteasome subunits and p97 ATPase subunits after knockdown of the Keap1 E3 ligase. Knockdown of Keap1 does not increase proteasome subunit levels.

Figure 5. The Cnc-C transcription factor contributes to maintaining basal levels of proteasome mRNAs and is essential for proteasome recovery after inhibition. For each quantitative PCR triplicate knockdown measurements for each target were carried out independently and the mRNA levels were determined using real time PCR and Taqman probes for *Drosophila* proteasome subunits and the p97 ATPase. A control RNAi treatment used dsRNA to the GFP sequence and was used for normalization of levels. A. The effects of knockdown of Cnc-C on a 20S proteasome subunit mRNA, alpha 7; on a 19S lid subunit, Rpt3; a 19S lid subunit, Rpn11; finally the mRNA levels for p97 ATPase. The Cnc was depleted with either a dsRNA that was specific for only the Cnc-C alternative transcript or for all Cnc transcripts, Cnc-All. Cnc-C RNAi depletion resulted in significant decreases in all measured proteasome mRNAs, Student's *t*-test value ($p < 0.0065$) for proteasome mRNAs and for p97 ATPase ($p = 0.037$). B. RNAi knockdown of the proteasome subunit S5a and increased levels of proteasome mRNA levels. Co-knockdowns of S5a and Cnc-C block the induction of proteasome and p97 mRNA levels. Depletion of Cnc-C with either dsRNA significantly blocks mRNA proteasome induction after loss of S5a. C. Proteasome recovery after inhibition with MG132 is blocked in cells previously RNAi depleted of the transcription factor Cnc-C. RNAi treated cells were treated with 7 hr MG132, 10 $\mu$M. Student's *t*-test value comparing control cell recovery after MG132 versus Cnc-C depleted MG132 treated cells, $p < 0.02$ for alpha 7 mRNA levels, $p < 0.007$ for p97 ATPase mRNA levels and $p < 0.0005$ for Rpn11.

Figure 6. A. Soluble *Drosophila* Cnc-C is a proteasome substrate and stabilized by RNAi depletion of proteasome subunit S5a. A C-terminal V5 epitope tag Cnc-C construct was transfected into *Drosophila* S2 cells and divided into different treatments. Collected cells were freeze-thawed and the

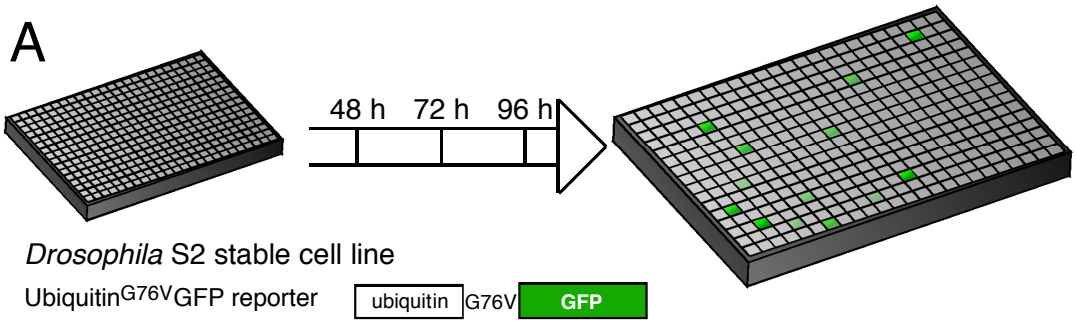
MCB Accepts published online ahead of print

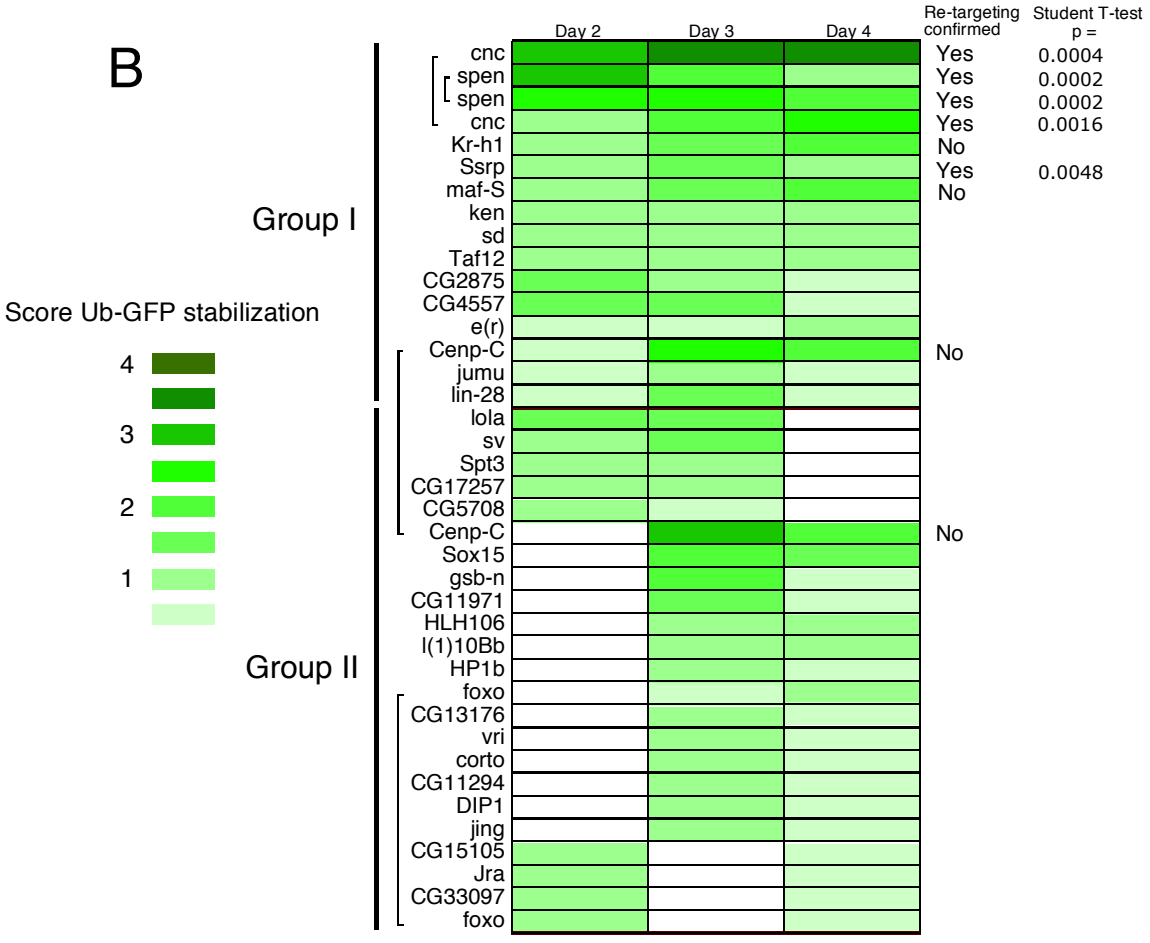
soluble protein fractions immunoblotted and developed with a V5 antibody. RNAi depletion of S5a stabilizes soluble Cnc-C but levels are not stablilized with a control dsRNA to CG10372. RNAi using dsRNA that was used in the original screen knocks down the CncC-V5 in epoxomicin treated cells. B. Soluble Cnc-C transfected *Drosophila* S2 cells expressing CncC-V5 were treated and soluble and membrane fractions immunoblotte d against the C-terminal V5 epitope. The faster migration for soluble Cnc-C suggests cleavage at the N-terminal end of the CncC-V5 compared to the slower migrating membrane associated CncC-V5. Both soluble and membrane Cnc-C bands migrate slower than the expected molecular weight for the Cnc-V5 construct, 66 kDa.

Figure 7. Speculative model for a *Drosophila* transcription feedback control mechanism for 26S proteasome levels. Unlike the mammalian system only a single locus exits in *Drosophila* for a cap 'n' collar (cnc) BZIP transcription factor. An alternative short transcript, Cnc-B is identical in its sequence to the longer Cnc-C transcript but the shorter form has been previously shown to transcribe a specific developmental gene class (51). Past and current evidence suggests a model where the Cnc-C is able to regulate the two gene classes independently and that post-transcriptional modifications of the Cnc-C protein could generate two forms of Cnc-C with distinct activities. The model proposes that a Cnc-C form that does not bind Keap1 functions similarly to the mammalian Nrf-1 protein and uses the Nrf-1 like N-terminal region to be membrane bound to the ER or other cell membrane surfaces. Depletion of the S5a shuttle protein activates this form of Cnc-C by preventing the degradation of the proteasome transcription factor. Regulation of oxidative stress genes is controlled by the level of a post-transcriptionally modified Cnc-C that functions like the mammalian Nrf-2 transcription factor and is controlled by poly-ubiquitylation through the Keap1 E3 ligase system. After entry into the nucleus the Cnc-C proteins binds additional proteins and functions as a transcription factor for either proteasome degradation or anti-oxidative stress class of genes.

MCB Accepts published online ahead of print

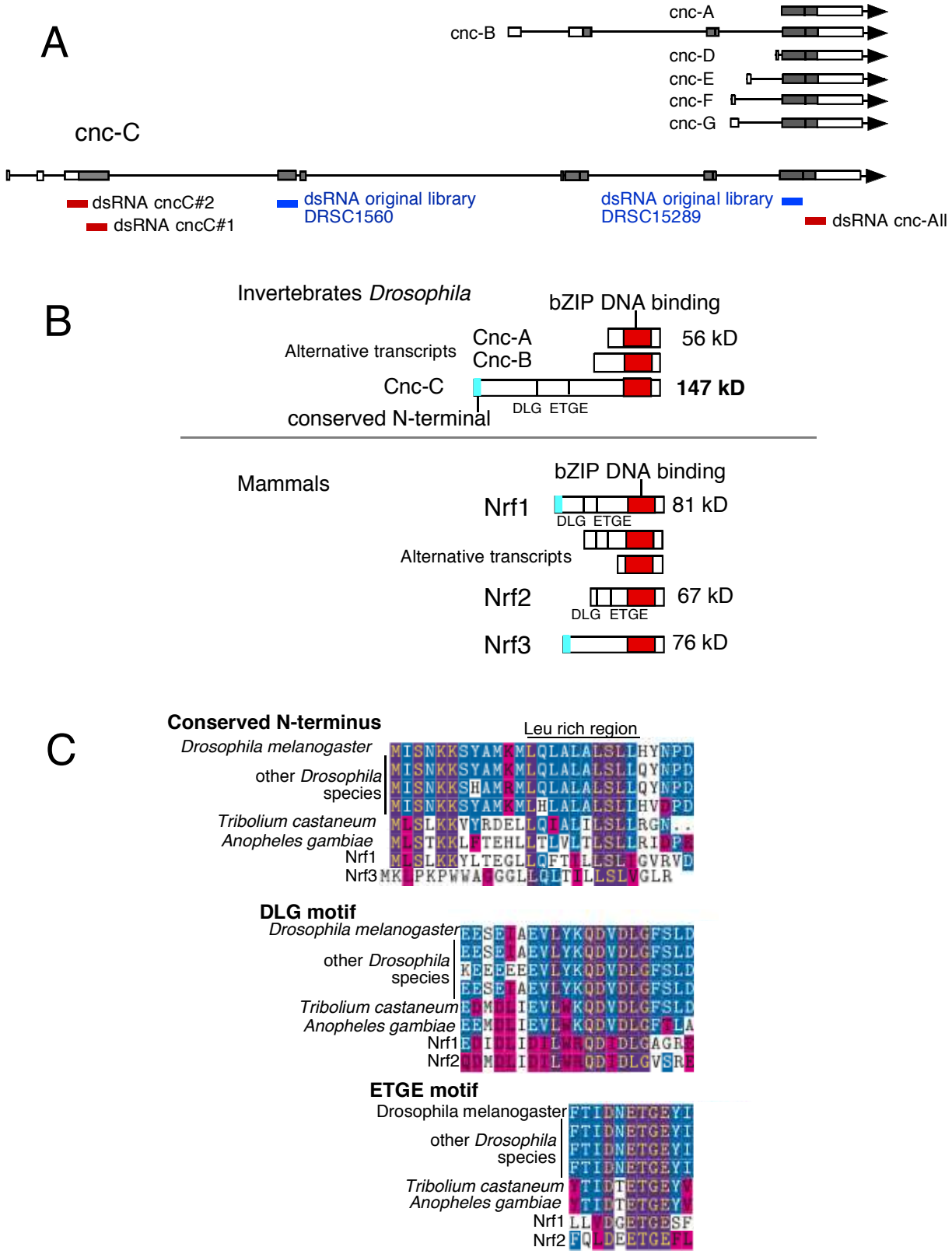
*Figure 1*



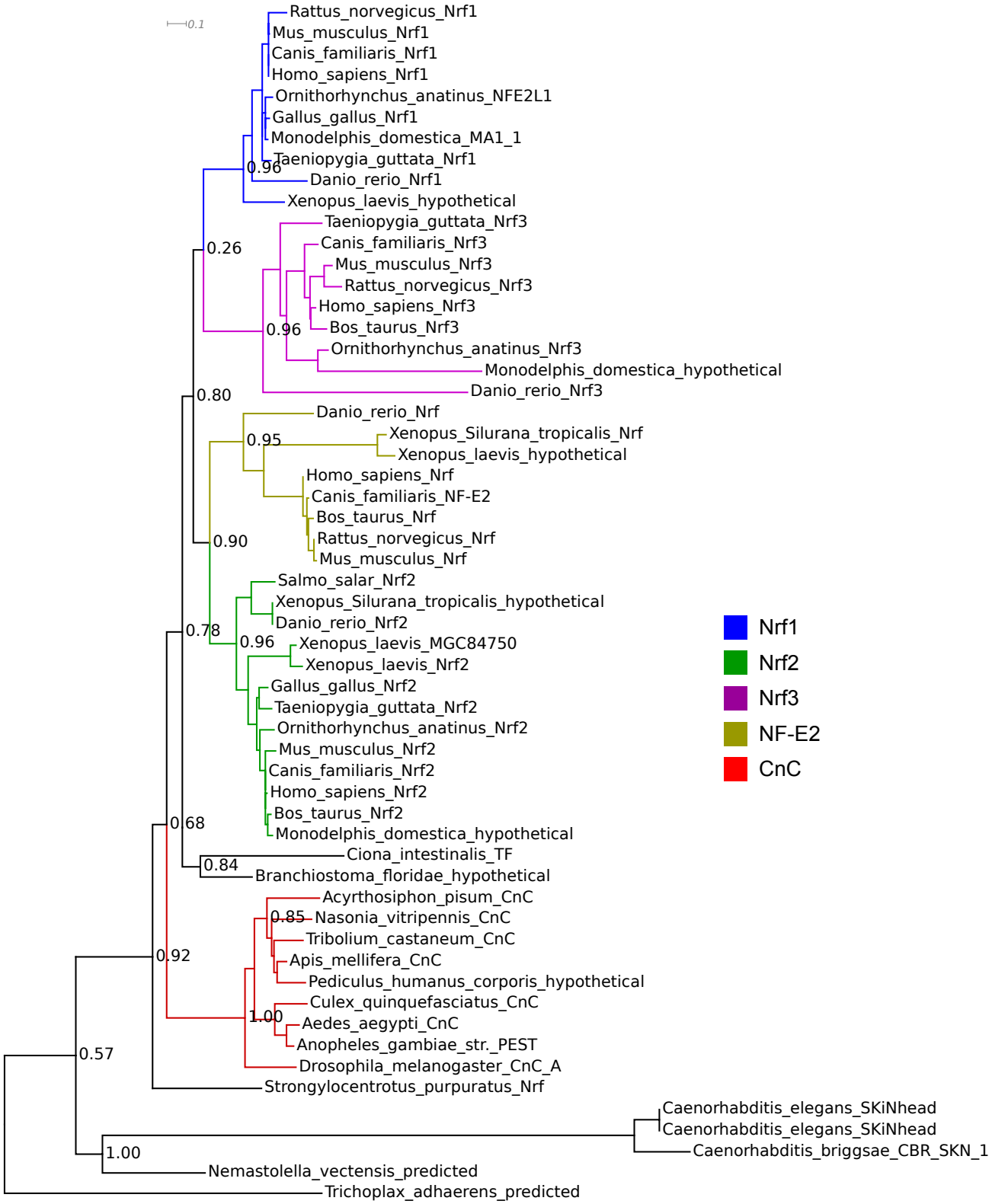



MCB Accepts published online ahead of print

*Figure 2*

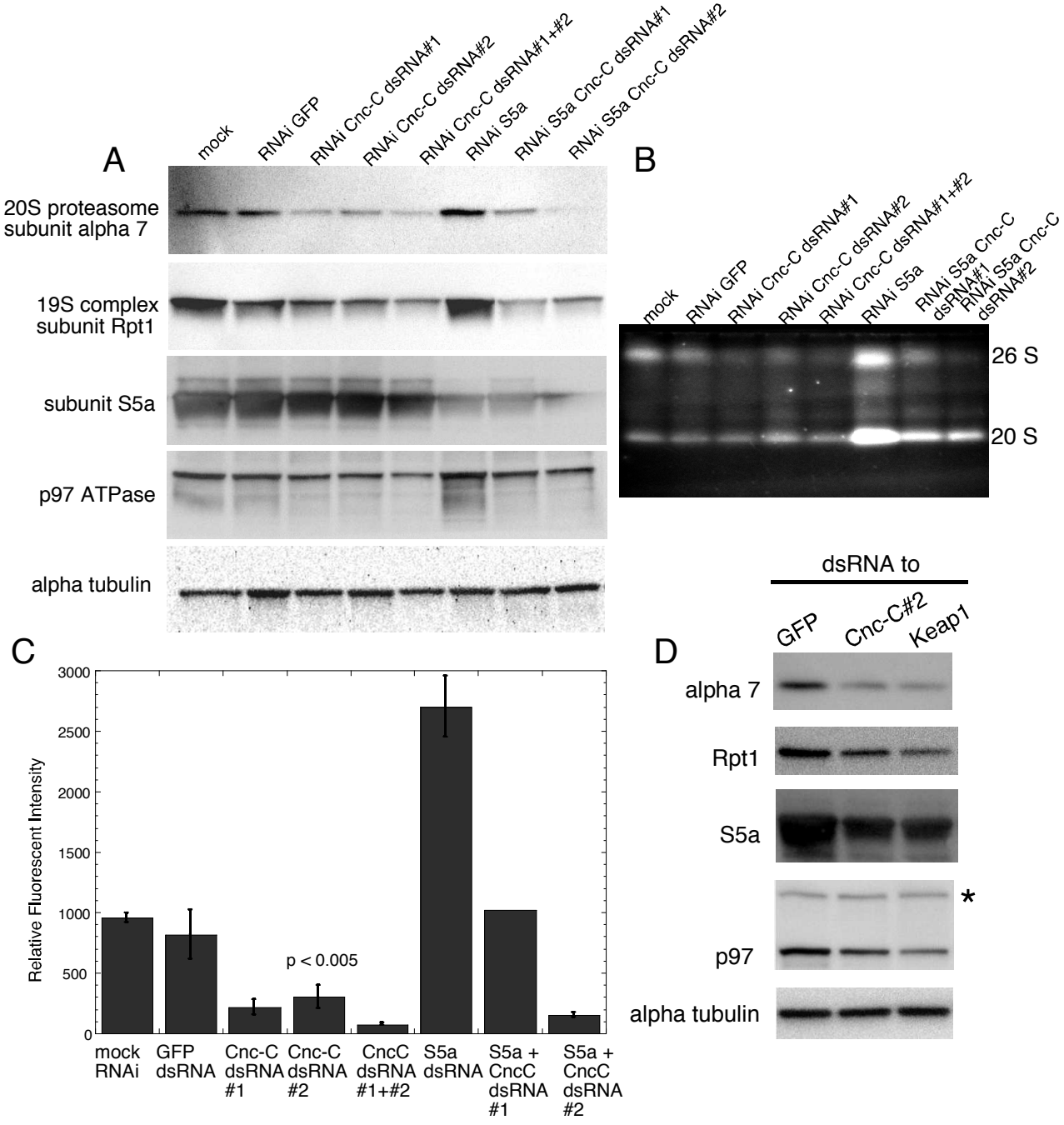


MCB Accepts published online ahead of print



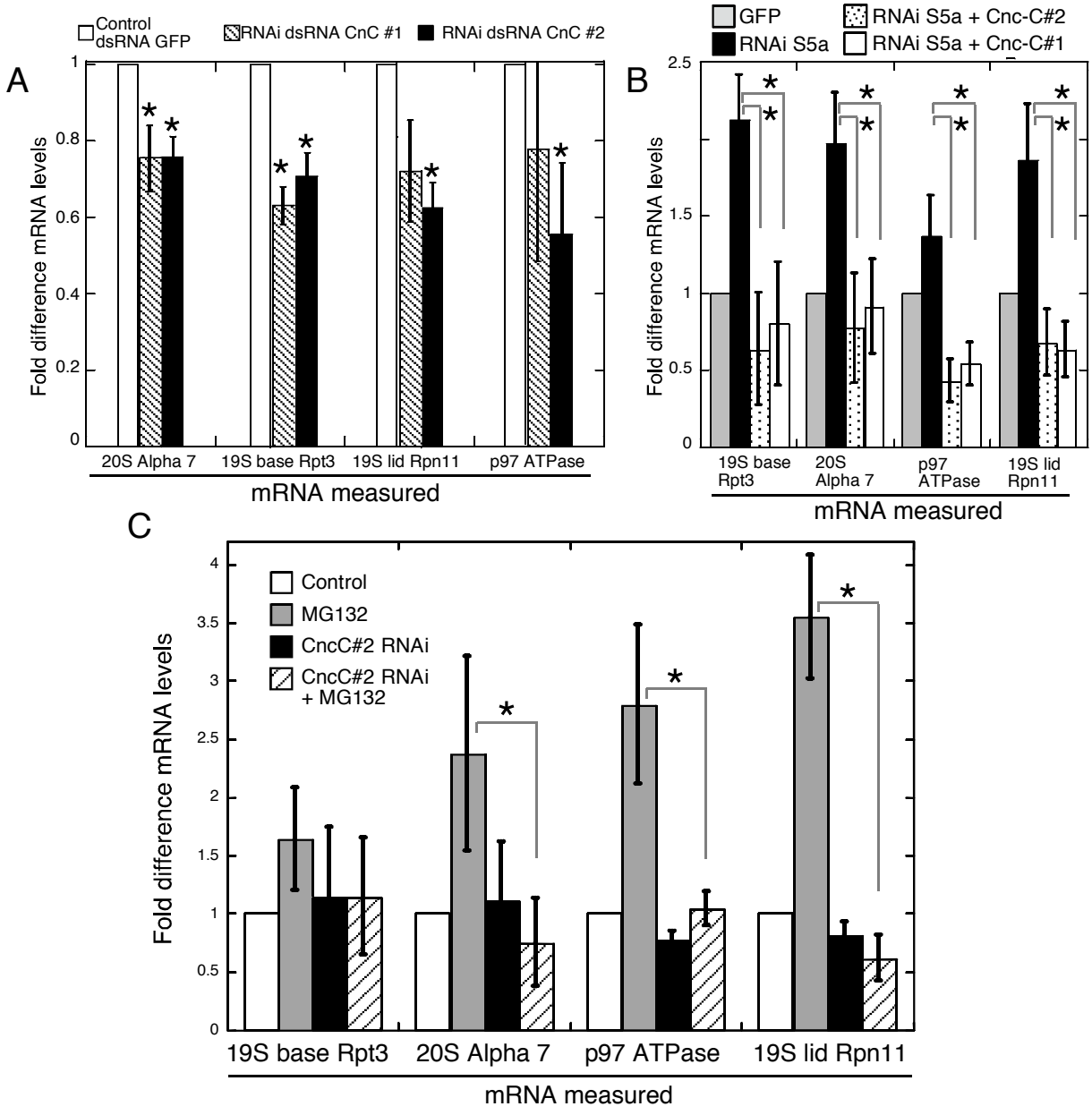
MCB Accepts published online ahead of print

*Figure 4*



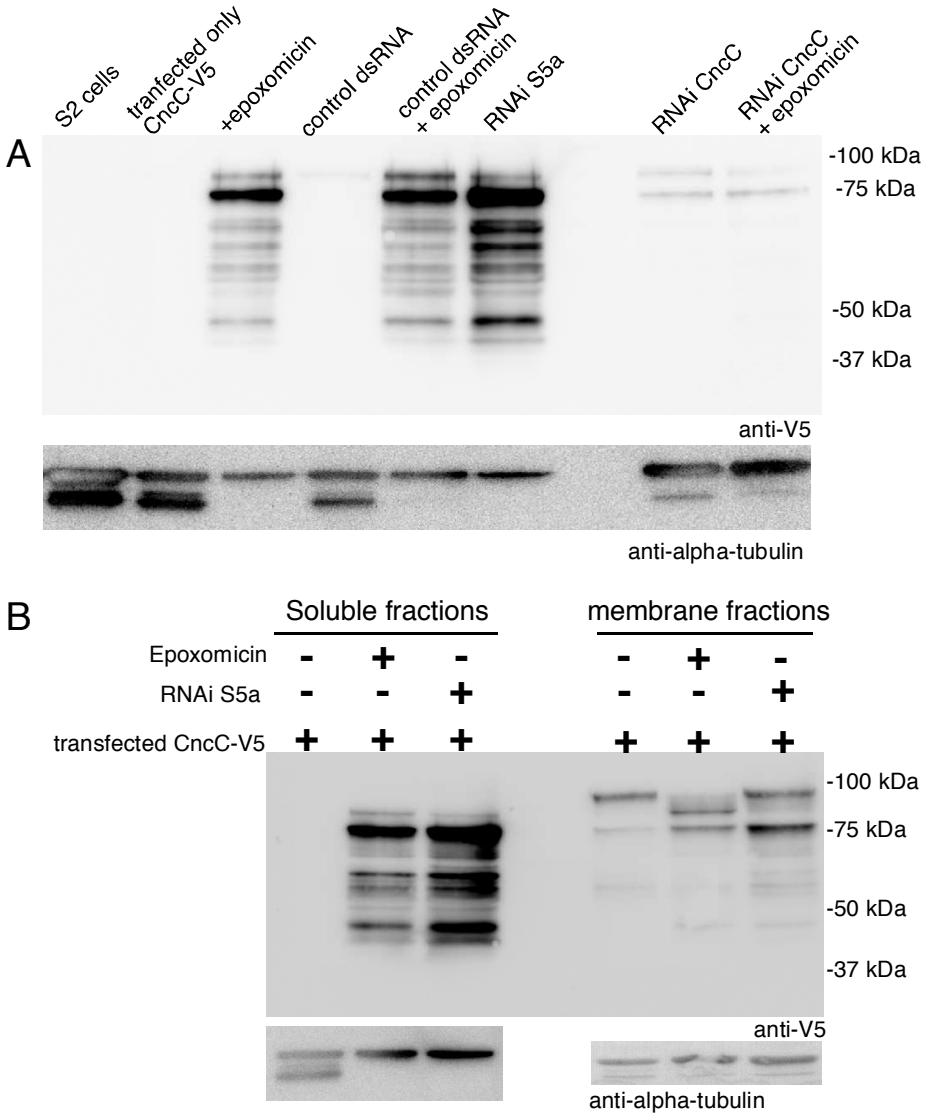
MCB Accepts published online ahead of print

*Figure 5*



MCB Accepts published online ahead of print

*Figure 6*



MCB Accepts published online ahead of print

*Figure 7*