Expression of the cytoplasmic nucleolin for post-transcriptional regulation of
macrophage colony-stimulating factor mRNA in ovarian and breast cancer cells

Ho-Hyung Woo*, Sang C. Lee, Steven J. Gibson, Setsuko K. Chambers

University of Arizona Cancer Center, Tucson, AZ 85724, USA

*Corresponding author at: University of Arizona Cancer Center, Tucson, AZ 85724, USA.

Email: hhwoo@email.arizona.edu_(H.H. Woo).

ABSTRACT

The formation of the mRNP complex is a critical component of translational regulation and mRNA decay. Both the 5' and 3'UTRs of CSF-1 mRNA are involved in post-transcriptional regulation. In CSF-1 mRNA, a small hairpin loop structure is predicted to form at the extreme 5' end (2-21nt) of the 5'UTR. Nucleolin binds the hairpin loop structure in the 5'UTR of CSF-1 mRNA and enhances translation, while removal of this hairpin loop nucleolin binding element dramatically represses translation. Thus in CSF-1 mRNA, the hairpin loop nucleolin binding element is critical for translational regulation. In addition, nucleolin interacts with the 3'UTR of CSF-1 mRNA and facilitates the miRISC formation which results in poly (A) tail shortening. The overexpression of nucleolin increases the association of CSF-1 mRNA containing short poly (A)_{n≤26}, with polyribosomes. Nucleolin both forms an mRNP complex with the eIF4G and CSF-1 mRNA, and is co-localized with the eIF4G in the cytoplasm further supporting nucleolin's role in translational regulation. The distinct foci formation of nucleolin in the cytoplasm of ovarian and breast cancer cells implicates the translational promoting role of nucleolin in these cancers.

Keywords:

Cytoplasmic nucleolin, Untranslated regions, Hairpin loop structure, eIF4G, Translation, Deadenylation, Macrophage colony stimulating factor (CSF-1) mRNA

1. Introduction

The untranslated regions (UTRs) in mRNA have a significant role in post-transcriptional regulation. Both the 5' and 3'UTRs are involved in mRNA stability and translation efficiency [1]. The regulatory elements found in the 5'UTR includes the 5'cap, secondary structure, and internal ribosome entry site (IRES). The 5'cap is a binding site for eIF4F, which is composed of the 5'cap-binding protein eIF4E, the scaffolding protein eIF4G, and the RNA helicase eIF4A1. The formation of the eIF4F complex recruits the 43S preinitiation complex, a major rate-limiting step in translation initiation [2]. The 5'cap also protects the mRNA from decay process [36].

The secondary structure in the 5'UTR modulates translation by binding proteins or other factors. The stem-loop structures in the 5'UTR can activate or inhibit translation depending on mRNA species and cellular circumstances [3,4]. A hairpin loop located near the 5'cap with a $\Delta G = -25 \sim -35$ kcal/mol has been described using the GFP reporter mRNA to inhibit translation initiation by blocking the access of the 43S preinitiation complex [5]. Dmitriev *et al* (2009) reported that the translation of mRNAs with long and highly structured 5'UTRs depends strongly on the 5'cap [3]. In contrast, highly structured IRES facilitates a cap-independent translation initiation [4,6].

The 3'UTR contains miRNA target sites, AU-rich elements (AREs), and poly (A)_n tail, and is involved in the determination of mRNA fate and translational regulation. Binding of miRNA forms miRISC in the 3'UTR, inducing mRNA deadenylation followed by degradation [7,8]. The poly (A)_n tail is also involved in mRNA stability and translation. Poly (A) binding protein (PABP) acts as a scaffold for many other proteins and regulates translation [9]. PABP interacts with eIF4F and mediates 5'-3' interaction, resulting in mRNP closed-loop formation

and facilitating the transfer of ribosomal subunits from the 3' - to the 5'-end for recycling [10,11,36,37].

Following translation, mRNA is assembled by the deadenylation complex, which leads to the decay process [7,12]. In humans, deadenylation is a major step in the decay process, which is followed by decapping and the exonuclease digestion of mRNA. Deadenylation is a biphasic process with consecutive actions of the Pan2-Pan3 and Ccr4-NOT deadenylation complexes. Pan2-Pan3 triggers the first phase of poly(A)_n tail shortening to ~110 nt, and Ccr4-NOT then follows by shortening the poly(A)₁₁₀ tail to 10-110 nt [13]. In miRNA-directed mRNA decay, the formation of the miRISC on the 3'UTR induces rapid mRNA decay by accelerating biphasic deadenylation [14]. Similarly, AU-rich elements (AREs) in the 3'UTR are destabilizing elements [15] which are involved in mRNA decay by triggering biphasic deadenylation [16].

Our recent findings, however, indicate that nucleolin mediates the miRNA-directed deadenylation of CSF-1 mRNA but increases translation without affecting the total mRNA level or half-life of the deadenylated form of CSF-1 mRNA [17]. This suggested that nucleolin-induced mRNA deadenylation may have a previously unidentified function in regulating translation. Nucleolin up-regulates the translation of a group of mRNAs by binding the structured (folded) region, *i.e.*, intra RNA-RNA hybrid structure [18], G-quadruplex [17], and AU-rich elements [19,20,21]. The CSF-1 mRNA, which encodes macrophage colony stimulating factor [22], a cytokine with roles in progression of several cancers including ovarian and breast cancers, interacts with nucleolin through the G-quadruplex and AREs in the 3'UTR [17,23]. The stability and translation of CSF-1 mRNA is also regulated by miRNAs [17,24]. Our findings indicated that the binding of nucleolin to the G-quadruplex and AREs in the CSF-1 mRNA

3'UTR not only mediates miRNA-directed CSF-1 mRNA deadenylation, but also blocks further miRISC-directed mRNA decay and increases the translation of CSF-1 mRNA [17].

In this paper, we find that nucleolin foci can be visualized in the cytoplasm of ovarian and breast cancer cells, and co-localizes with eIF4G. Overexpression of nucleolin in the cytoplasm increases the polyribosome-associated CSF-1 mRNA containing shortened poly (A)_{n≤26} tails. A hairpin loop structure is predicted at the extreme 5' end (2-21nt) of the 5'UTR. Nucleolin interacts with this hairpin loop structure in the 5'UTR, increasing translation only in the presence of this hairpin loop nucleolin binding element. Removal of this element encoding the hairpin loop structure in the 5'UTR dramatically represses translation of mRNA. Collectively, our data suggests that the interaction of nucleolin with eIF4G and with the hairpin loop structure in the 5'UTR of CSF-1 mRNA are largely responsible for enhanced CSF-1 mRNA translation and stability.

2. Materials and methods

2.1. Cell lines

Hey and SKOV3 ovarian cancer cell lines were cultured in DMEM/F12 (Mediatech) supplemented with 10% fetal bovine serum. BT20 breast cancer cell line was cultured in MEM supplemented with 10% fetal bovine serum. MDA-MB-231 breast cancer cell line was cultured in DMEM supplemented with 10% fetal bovine serum. Hey ovarian cancer cell lines with chromosomally-integrated nucleolin (NCL)-GFP or GFP were generated.

2.2. Preparation of cytoplasmic and nuclear protein

Cytoplasmic and nuclear protein fractions were isolated using NE-PER Nuclear and Cytoplasmic Extraction Kit (Pierce).

2.3. Indirect immunofluorescence microscopy

The procedure described by Kedersha and Anderson [27] was followed. Cells were fixed in 4% paraformaldehyde for 5 min, and permeabilized in cold methanol for 15-20 min before incubation with the nucleolin antibody for 1 hr at 37°C. The N-terminal or C-terminal immunogen-raised rabbit monoclonal nucleolin antibodies (Abcam) and middle domain raised mouse monoclonal nucleolin antibody (Santa Cruz Biotechnology) were used. The Alexa-conjugated secondary antibodies (Life Technologies) were incubated for 1 hr at 37°C, and the resulting image was photographed under the fluorescence microscope.

2.4. Nucleolin shRNA transfection assay

Transfection and sequence information of nucleolin shRNA was described previously [17].

2.5. Luciferase reporter constructs

CSF-1 mRNA 5' and 3'UTRs were ligated with Firefly luciferase gene and cloned into the pTRE3G-BI plasmid (Clontech). Triple mutation was described previously [17].

For constitutive expression of reporter constructs, CSF-1 mRNA 5' and 3'UTRs were ligated with Firefly luciferase gene and cloned into the pcDNA3.1 plasmid.

2.6. Translation induction assay and ribosome profile

The 'Tet-ON' system provides better control of transcription as well as translation, since there is no accumulation of translation products (*i.e.*, luciferase) in the cell before the addition of tetracycline. Chromosomally-integrated SKOV3 and Hey cell lines with a Tet-ON 3G inducible system (Clontech) were generated with the plasmid constructs indicated in Fig. 2A.

In translation induction assay, cells were treated after 24 hr incubation for 4 hr with 20 ng/ml doxycycline which was then removed by PBS wash, 5 times. After doxycycline removal, cells were harvested and lysed every hour for the assay of luciferase activity. A ribosome profile was generated as described previously [17]. N=10.

2.7. RNA electrophoretic mobility shift assay

REMSA was described previously [26]. The sequences at extreme 5' end of the 5'UTR (2-21nt), 5'UTR without hairpin loop sequence (22-105nt), and G-quadruplex from the 3'UTR (2855-2892nt) were used to generate ³²P-labeled RNA probe for REMSA. ³²P-labeled RNA probe was solubilized in 50 mM Tris-HCl_{pH7.4} and 100 mM KCl to maintain secondary structure. Linear 5' hairpin RNA sequence is induced by solubilizing in 50 mM Tris-HCl_{pH7.4} without 100

mM KCl and heat denatured for 5 min before REMSA. Myc-DDK-tagged nucleolin (Origene) was purchased.

2.8. Analysis of luciferase RNA half-life

To determine luciferase reporter mRNA half-life in Hey ovarian cancer cells, actinomycin-D (Act-D) chase experiments were performed with 3 µg/ml of Act D (Sigma) added to inhibit new transcription. For cells undergoing transfection of various constructs, the experiments were performed 2 days after transfection. Cells were harvested at 0 h, 1 h, 2 h, and 4 h after Act D treatment, total cellular RNA extracted using Trizol (Invitrogen), and luciferase RNA, GAPDH, or *Renilla* mRNA levels were analyzed by qRT-PCR. Luciferase mRNA half-lives were calculated after qRT-PCR, normalized to GAPDH mRNA, values were plotted, and the time period required for a given transcript to decrease to one-half of the initial abundance was calculated. GAPDH mRNA is not affected by nucleolin, and has a long half-life (>18 hrs). Three independent experiments were performed.

2.9. Native mRNP Immunoprecipitation

Native co-immunoprecipitation (Co-IP) of the endogenous mRNP complex was done by the protocol described previously with modifications [17,26]. Cytoplasmic proteins were extracted using NE-PER Nuclear and Cytoplasmic Extraction Reagents (Pierce). To maintain the intact native state of the mRNP complex, cytoplasmic lysates were not fixed or sonicated. To partially disrupt the mRNP complex, cytoplasmic lysates were treated with RNase A (10 ng/ul) and T1 (1 U/ul) for 15 min at 37°C before immunoprecipitation. For nucleolin IP, 5 µg mouse monoclonal anti-human nucleolin antibody raised against middle domain (Santa Cruz

Biotechnology), or rabbit monoclonal anti-human nucleolin antibodies raised against N- or C-terminal residues (Abcam), were used. For eIF4G IP, 5 µg rabbit polyclonal anti-human eIF4G antibody (Santa Cruz Biotechnology) was used. A reaction containing normal mouse or rabbit IgG (Sigma) served as a negative control. For IP-RT-qPCR, cells were sonicated and RNases were not added in IP.

2.10. Determination of deadenylated mRNAs associated with polyribosomes

Polyribosome profiling was described previously [17]. Hey cells were grown ~80% confluence in four 10 cm plates. After cycloheximide treatment for 15 min, cells were harvested, lysed, and fractionated by sucrose density gradient ultracentrifugation. RNA in each fraction was isolated by Trizol. RNA was divided equally. The first half of RNA was primed by pdN₆ and quantified by RT-qPCR to get the amounts of total mRNAs associated with polyribosomes. The second half of RNA was annealed and ligated with 5'-phosphorylated dT₁₇ for 30 min followed priming by dT₁₀-Adapter₁₄ for 3 h for LM-PAT analysis [25] (see Additional data file Figure S7). For quantification of long poly (A)_{n≥27} mRNAs in each fraction, RT-qPCR was done by sequence specific forward primer and dT₁₀-Adapter₁₄ primer. For analysis of the relative distributions of CSF-1, AKT1, and GAPDH mRNAs in polyribosome gradients, C_t values from individual fractions 1-10 were each subtracted from the C_t value from fraction 1 for CSF-1 and GAPDH mRNAs, as fraction 1 had the largest C_t values (that is, the lowest CSF-1, AKT1, and GAPDH mRNA abundances). The resulting ΔC_t numbers were converted into fold differences. The abundance of each mRNA as a percentage of the total from all 10 fractions was then calculated. Because GAPDH mRNA is not a binding target of nucleolin [17], its abundance was used as a control for estimation of relative distribution of CSF-1 mRNA in polyribosome gradients. N=6.

2.11. Sequencing of LM-PAT products

The LM-PAT products of CSF-1 mRNA are sequenced by sequence-specific nested primer (5'-agaggacattggctcacgcactgtg-3'). Sequencing reaction is done three times from independent samples.

2.12. Statistical Analyses

Data are depicted as mean $\pm$ SD from at least three independent experiments. Exact *n* values are provided in the figure legends. The unpaired two-way *t*-test and one-way ANOVA test were performed using SigmaStat (Jandel Scientific Corp.). *P*<0.05 was considered statistically significant.

3. Results

3.1 Formation of distinct foci of nucleolin in the cytoplasm of breast and ovarian cancer cells

Immunoblot analyses using nucleolin antibodies raised against immunogens from either the N-terminal acidic region, middle domain for RNA binding, or C-terminal RGG repeats (Fig. 1A) detect the same characteristic nucleolin in both cytoplasmic and nuclear fractions of Hey ovarian cancer cells (Fig. 1B), *i.e.*, the cytoplasmic form of nucleolin (cyto-NCL) has the same characteristics as the nuclear form of nucleolin (nu-NCL) in both SDS-PAGE (~110 kDa) (Fig. 1B) and IEF ($pI \cong 4.6$) (Fig. 1D). The immunoblot analyses in the cytoplasmic fraction of Hey ovarian cancer cells, in which nucleolin is downregulated by shRNA, confirm the specificities of nucleolin antibodies (Supplementary data Figure S1).

Previously, subcellular localization studies by immunofluorescence (IF) microscopy indicate that nucleolin is primarily localized in the nucleus using the middle domain raised antibody [28]. Since the cytoplasmic form of nucleolin (cyto-NCL) is detectable by immunoblot analysis with all 3 antibodies (Fig. 1B), the cyto-NCL should be detectable using indirect IF microscopy. We examined MCF10A non-cancerous breast epithelial cells, two breast cancer cell lines (BT20 and MDA-MB-231), and two ovarian cancer cell lines (Hey and SKOV3). The immunoblot analysis showed similar expression level in five different cell lines in both cytoplasmic and nuclear fractions (Supplementary data Figure S2). However, in non-cancerous breast epithelial MCF10A cells, IF microscopy shows that nucleolin is primarily detected in the nucleus (Fig. 1E), *i.e.*, the N-terminal raised antibody did not detect nucleolin in the nucleus and cytoplasm or the signal was very faint if any, and the middle-domain and C-terminal raised antibodies detected nucleolin primarily in the nucleus, with faint signal seen in the cytoplasm only with use of the middle domain raised antibody. In contrast, using breast and ovarian cancer

cell lines, the N-terminal raised antibody detected mainly the cytoplasmic form of nucleolin with minor signals in the nucleus (MDA-MB-231, Fig. 1F; BT20, Supplementary data Figure S3; Hey, Figure S4, and SKOV3, Figure S5). Thus, although immunoblot showed both cytoplasmic and nuclear forms of nucleolin, IF microscopy using the N-terminal raised antibody could only visualize the cytoplasmic form of nucleolin in breast and ovarian cancer cells. In contrast, the middle domain raised antibody detected both the cytoplasmic and nuclear forms of nucleolin in breast cancer cells (MDA-MB-231 in Fig. 1F and BT20 in Supplementary data Figure S3). In ovarian cancer cells, the middle domain raised antibody detected the nuclear form of nucleolin (Hey and SKOV3 in Supplementary data Figures S4 and S5). Depending on the cell line, the strength of the cyto-NCL signal detected by the middle domain raised antibody varied; it was strong in MDA-MB-231, and BT20, but very weak in MCF10A and not detectable in Hey and SKOV3. Results using the C-terminal raised antibody differed from that of the 2 previous antibodies in that it localized nucleolin mainly in the nucleus of MCF10A and four different cancer cell lines (Fig. 1E and F, Supplementary data Figures S3, S4, and S5). When using the C-terminal raised antibody, despite visualization in the cytoplasmic form of nucleolin by immunoblot, the cytoplasmic form could not be clearly visualized by IF microscopy.

To confirm cyto-NCL in the cytoplasm, we down-regulated nucleolin by shRNA. Down-regulation of nucleolin decreased the cyto-NCL signal in the cytoplasm of MDA-MB-231 cells (N-terminal (shRNA) in Fig. 1F), indicating the cyto-NCL signal detected is nucleolin. Furthermore, the cyto-NCL detected by the N-terminal raised antibody overlapped with the cyto-NCL detected by the middle domain raised antibody, indicating that the cyto-NCL detected by both antibodies is the same protein (Fig. 1G). Since immunoblot analysis shows similar expression level of cyto-NCL in all four cancer cell lines (Supplementary data Figure S2), taken

together, we conclude that cyto-NCL forms distinctive foci in the cytoplasm of both ovarian and breast cancer cells.

3.2 The extreme 5' end of the 5'UTR in CSF-1 mRNA is responsible for translational regulation

We previously reported that nucleolin interacts with the G-quadruplex (2855-2892nt) and AREs (3796-3900nt) in the 3'UTR of CSF-1 mRNA [17]. The binding of nucleolin to the G-quadruplex and AREs facilitates the miRISC formation to the nearby miRNA target sequence (2571-2580nt), resulting in deadenylation. Mutation of three *cis*-acting elements (G-quadruplex, miRNA target, and ARE elements) in 3'UTR of CSF-1 mRNA dramatically increases the luciferase reporter mRNA level [17].

In the present report, we studied the role of the 5'UTR of CSF-1 mRNA in post-transcriptional regulation. The 5'UTR of CSF-1 mRNA is GC-rich (76.4%). Bioinformatics analysis indicates that the extreme 5' end (2-21nt) of the 5'UTR of CSF-1 mRNA can form a weak secondary structure ($dG = -7.7$ kcal/mol), a small hairpin loop (hl) (Supplementary data Figure S6). To test whether this sequence is involved in post-transcriptional regulation, four constructs with luciferase gene fused with the CSF-1 mRNA 5' and 3'UTRs under the control of a tetracycline-inducible promoter were made, *i.e.*, Wt=Wt (wild type construct), D(21)=Wt (deletion of 21nt in the extreme 5' end), Wt=Mut(Trip1) (Mutation in the G-quadruplex, miRNA target, and ARE elements in the 3'UTR; detailed description in ref. 17), and D(21)=Mut(Trip1) (mutation both in the 5' and 3'UTRs) (Fig. 2A). In the wild-type construct (Wt=Wt), both the 5' and 3'UTRs have the complete sequence of the CSF-1 mRNA UTRs, including the hl in the 5'UTR and miRNA-target, G-quadruplex, and AREs in the 3'UTR.

To examine the role of these *cis*-acting elements in the CSF-1 mRNA UTRs on post-transcriptional regulation, we generated three different chromosomally-integrated mutant constructs and a wild-type construct in SKOV3 and Hey ovarian cancer cell lines (Fig. 2A). The overall pattern of luciferase activities between the SKOV3 and Hey cell lines was similar, with only minor differences (Fig. 2B and 2C). In all constructs, luciferase activity increased linearly (Fig. 2B and 2C). The ratio of luciferase activity divided by induction time (L/t) indicates that the L/t reached the plateau by 3-5 hr induction (Fig. 2D and 2E).

Effects of the G-quadruplex, miRNA target, and AREs in 3'UTR: The most dramatic effect was observed in Wt=Mut(Tripl) which is a mutation of the miRNA target, G-quadruplex, and ARE sequences in the 3'UTR, in which by 5 hr induction luciferase activity had increased 5.3-fold ($P<0.001$) in SKOV3 cells and 12.5-fold ($P=0.001$) in Hey cells when compared to the Wt=Wt (Fig. 2B and 2C). We previously reported that by using a constitutive expression system (under the control of the CMV promoter), the triple mutation dramatically increased luciferase activity and mRNA level when compared to the wild-type construct in the steady-state condition [17]. We also showed that individual mutations of *cis*-acting elements do not have as much of an effect when compared to the triple mutation construct [17]. From this experiment, we conclude that all three *cis*-acting elements work together to repress CSF-1 mRNA translation in a time-dependent manner. However, our prior report [17] had uncovered the complexities and double-edged sword effect of nucleolin leading to deadenylation of CSF-1 mRNA, while promoting translation.

Effects of 5'UTR sequences encoding the hl: To test the role of the 5'UTR small hl, two mutant constructs {D(21)=Wt, D(21)=Mut(Tripl)} were tested (Fig. 2A). The deletion of the hl sequence decreased luciferase activity 2.47-fold ($P<0.001$) after 5 hr when compared to the

Wt=Wt in SKOV3 ($L/t=1.08$ at 5 hr in D(21)=Wt, $L/t=2.67$ at 5 hr in Wt=Wt). (Fig. 2B and 2D).

The most dramatic effects were observed when the hl was deleted in the triple mutant. The mutation of the hl in the 5'UTR of the triple mutant {D(21)=Mut(Tripl)} decreased luciferase activity dramatically ($L/t=1.07$ at 5 hr) compared to the Wt=Mut(Tripl) ($L/t=14.13$ at 5 hr) ($P=0.045$) (Fig. 2B and 2D). Similar effects were also observed in Hey cells (Fig. 2C and 2E).

Effects on mRNA level: However, importantly, there were no significant differences in total mRNA levels between the Wt=Mut(Tripl) and D(21)=Mut(Tripl) constructs in both SKOV3 ($P=0.819$) and Hey cells lines ($P=0.218$) (purple bar in Fig. 2F and G). In both constructs, the mRNA levels were 4~5-fold higher than that of the Wt=Wt construct. In contrast, total mRNA levels in D(21)=Wt construct decreased significantly compared to the Wt=Wt construct, *i.e.*, ~92.3% ($P=0.0343$) in SKOV3 and ~60.7% ($P=0.038$) decreased in Hey cells, respectively.

Together, our data indicates that three *cis*-acting elements in the 3'UTR work together to downregulate translation by inducing miRNA-directed deadenylation followed by decay of CSF-1 mRNA [17]. Their ability to induce mRNA decay appears to be critically dependent on the presence of the miRNA target, G-quadruplex, and AREs in the 3'UTR [17]. In contrast, the hl in the 5'UTR has a strong translation-enhancing effect. The hl in the 5'UTR also appears to stabilize the CSF-1 mRNA, since the removal of the sequences encoding the hl decreased luciferase mRNA level significantly.

3.3. Nucleolin interacts with the extreme 5' end of the 5'UTR in CSF-1 mRNA and increases translation

Since nucleolin increases translation and mRNA stability [17] and the 5'UTR element encoding the hl structure is critical for translational regulation (Fig. 2), we tested the interaction

of nucleolin with the hl structure of the 5'UTR. RNA electrophoretic mobility shift assay indicates that nucleolin interacts with the hl structure RNA from the extreme 5'end of the 5'UTR, as well as with the G quadruplex RNA from the 3'UTR of CSF-1 mRNA, a known nucleolin binding structure (Fig. 3A) [17]. In contrast, nucleolin does not bind either the linear hl sequence or the 5'UTR sequence without hl. The REMSA for the structured 5'hl was performed under conditions which preserved secondary structure, thus this data suggests binding of nucleolin to the hl structure.

To confirm nucleolin interaction with the hl structure at the 5'end of 5' UTR, we tested whether the wild-type construct (Wt=Wt) and the hl deletion construct {D(21)=Wt} respond to the overexpression of nucleolin. The expression of both constructs is under the control of CMV promoter for constitutive expression. We confirmed that the overexpression of nucleolin increased the luciferase protein level in wild type construct containing both the 5'UTR hl structure, as well as the three *cis*-acting elements in the 3'UTR (Fig. 3B). In contrast, the hl deletion construct, D(21)=Wt, did not respond to the nucleolin overexpression.

We previously reported that overexpression of nucleolin enhances deadenylation of CSF-1 mRNA and translation [17]. In Wt=Wt construct, the half-life of luciferase RNA was ~2.3h with or without nucleolin overexpression (Fig. 3C). In contrast, D(21)=Wt construct the half-life was ~1 h with or without nucleolin overexpression. The nucleolin overexpression did not influence the mRNA stability in both Wt=Wt (P=0.905) and D(21)=Wt (P=0.686) constructs.

We conclude that the interaction of nucleolin with the hl structure enhances translation of CSF-1 mRNA. In the presence of the hl structure in 5'UTR, nucleolin appears to be able to interfere with the mRNA decay which generally follows the miRNA-directed deadenylation, mediated by the three *cis*-acting elements in the 3'UTR of CSF-1 mRNA [17].

3.4 Nucleolin interacts with eIF4G to form an mRNP complex in the cytoplasm

Since translation initiation is a rate-limiting step and nucleolin overexpression increases the translation of CSF-1 mRNA (Fig. 3B), the interaction of nucleolin with eIF4G in the translation initiation complex was investigated using native co-immunoprecipitation assay (co-IP). Since the translation initiation complex, together with nucleolin, can form a large multicomponent complex, the mRNP complex needs to be partially relaxed for antibody access. Samples were either RNase A and T1 treated to relax mRNA complexes or untreated. In the context that cytoplasmic nucleolin was strongly detected by the N-terminal antigen raised antibody in IF microscopy (Fig. 1), we tested all three types of nucleolin antibodies, *i.e.*, N-terminal (N-NCL), middle domain (M-NCL), or C-terminal (C-NCL) antigen raised antibodies in Hey ovarian cancer cells. We found that the N-NCL antibody immunoprecipitates nucleolin only in RNases-treated sample, in which nucleolin co-immunoprecipitates with eIF4G (Fig. 4A). Without RNase-treatment, nucleolin did not immunoprecipitate nor did it co-immunoprecipitate with eIF4G, suggesting a large translation initiation complex together with nucleolin needs to be partially relaxed for successful co-IP with this antibody. The M-NCL antibody immunoprecipitates nucleolin in all treatments, but nucleolin did not co-immunoprecipitate with eIF4G (Fig. 4B). Similarly, the C-NCL antibody immunoprecipitates nucleolin, but nucleolin did not co-immunoprecipitate with eIF4G (Fig. 4C).

In contrast, the eIF4G antibody immunoprecipitates eIF4G in all treatments (Fig. 4D). However, eIF4G co-immunoprecipitates with nucleolin only in RNases-treated sample.

To confirm nucleolin-eIF4G mRNP complex contains CSF-1 mRNA, RT-qPCR was performed. mRNP complex immunoprecipitated by either nucleolin or eIF4G antibody contains CSF-1 mRNA (Fig. 4E).

We conclude that nucleolin forms an mRNP complex with eIF4G for translational regulation. For co-IP assay, this mRNP complex needs to be partially relaxed for nucleolin antibody access. We propose that nucleolin serves as a core in the mRNP complex, *i.e.*, both middle domain and C-terminal of nucleolin are occupied in the mRNP complex and therefore the N-terminal is only available for antibody access to enable co-IP of eIF4G. The free-form of nucleolin is immunoprecipitated in cytoplasmic lysates by M- and C-terminal antigen raised antibodies (Fig. 4). This free-form of nucleolin does not interact with eIF4G and does not form an eIF4G containing mRNP complex (Fig. 4).

3.5. Cytoplasmic form of nucleolin co-localizes with eIF4G in cell

Since co-IP data indicates that nucleolin interacts with eIF4G (Fig. 4), the co-localization of nucleolin with eIF4G *in cell* was observed by indirect IF microscopy. The cyto-NCL was co-localized with eIF4G in the cytoplasm of MDA-MB-231 breast cancer cells (Fig. 5).

3.6. Nucleolin overexpression shortens poly(A)_n tails of CSF-1 mRNA associated with polyribosomes in the cytoplasm

Since nucleolin forms distinct foci in the cytoplasm of ovarian and breast cancer cells (Fig. 1) and increases translation of CSF-1 mRNA with short poly (A) tail [17], we performed a detailed analysis of the CSF-1 mRNA status in the nucleolin overexpressed Hey ovarian cancer cells. We previously reported that nucleolin mediates miRNA-directed deadenylation, but increases translation of CSF-1 mRNA whose 3'UTR interacts with miRNAs [17,24]. To check the status of poly(A)_n tail length of CSF-1 mRNA associated with polyribosomes, we took advantage of a novel LM-PAT approach we designed using mRNAs fractionated by a sucrose density gradient (Supplementary data Figure S7). Since binding of PABP requires the poly(A)

tail length of ~27nt [29] and translation is coupled tightly with the poly(A) tail length of ~20nt [30], we designed the LM-PAT analysis to determine the poly(A) tail length <27nt in CSF-1 mRNA. Both chromosomally-integrated nucleolin-GFP and GFP Hey ovarian cancer cell cytoplasmic lysates were fractionated. The ribosome profiles were similar between Hey (NCL-GFP) and Hey (GFP) cells (Fig. 6A). The ribosome profile indicates that nucleolin is present in free ribosome fractions as well as in 40S, 80S, and polyribosome fractions (Fig. 6B). The distribution of the eIF4G overlaps with the distribution of nucleolin in the ribosome profile. Ribosomal proteins (Rps6 and RpL19) are present in both light and heavy polyribosome fractions.

To validate our method, the deadenylation status of GAPDH mRNA in the ribosome profile was tested. GAPDH mRNA was chosen as a negative control for nucleolin regulation, where the percentage of the deadenylated form is small under both control and nucleolin-regulated conditions. The total RNA from each fraction was isolated and used for cDNA synthesis. Half of the total RNA was primed by a random primer (pdN₆) for cDNA synthesis and amplified by qPCR to obtain the percent of total mRNA associated with polyribosomes. The other half of the RNA was annealed and ligated with phosphorylated dT₁₇ which followed priming by an oligo-dT₁₀-Adaptor₁₄ and was used for LM-PAT-RT-qPCR to obtain the percent of long poly(A)_{n≥27} mRNAs associated with polyribosomes. The percent of deadenylated mRNAs (*i.e.*, short poly(A)_{n≤26} mRNAs) associated with polyribosomes was obtained by subtracting the percent of long poly(A)_{n≥27} mRNAs associated with the corresponding sucrose density gradient fractions of polyribosomes from the percent of total mRNAs. Neither GAPDH protein expression nor the percent of deadenylated GAPDH mRNA associated with polyribosomes was affected by nucleolin overexpression (Supplementary data Figure S8). Of the

total GAPDH mRNA in control Hey cells (GFP) and in nucleolin-overexpressed Hey cells (NCL-GFP), $41.24 \pm 0.9\%$ and $38.91 \pm 1.53\%$, respectively, were associated with polyribosomes (fractions 6-10) (Supplementary data Figure S8). In control Hey cells (GFP), $35.59 \pm 1.34\%$ of GAPDH mRNA was long poly(A)_{n $\geq$ 27} mRNA associated with polyribosomes (fractions 6-10). In nucleolin-overexpressed Hey cells (NCL-GFP), $34.17 \pm 1.19\%$ of GAPDH mRNA was long poly(A)_{n $\geq$ 27} mRNA associated with polyribosomes (fractions 6-10). The percent of deadenylated GAPDH mRNA (*i.e.*, short poly(A)_{n $\leq$ 26} mRNAs) associated with polyribosomes was obtained by subtracting the percent of GAPDH total mRNA by the percent of GAPDH long poly(A)_{n $\geq$ 27} mRNA. In control Hey cells, 5.66% of deadenylated GAPDH mRNA (*i.e.*, short poly(A)_{n $\leq$ 26} mRNAs) was associated with polyribosomes, while 4.74% of deadenylated GAPDH mRNA was associated in nucleolin-overexpressed (NCL-GFP) Hey cells (fractions 6-10). There was no difference in either total or deadenylated GAPDH mRNA levels between two cell lines ($P=0.0854$ for total mRNAs; $P=0.242$ for long poly(A)_{n $\geq$ 27} mRNAs). These results are expected from the study of GAPDH mRNA.

Using this validated method, CSF-1 mRNA whose translation was influenced by nucleolin [17, 38] were then tested. CSF-1 protein level was increased in nucleolin-overexpressed Hey cells (Additional data file Figure S9A). However, there was no significant difference in the total mRNA levels of CSF-1 between nucleolin-overexpressed and control Hey cell lines ($P=0.821$ in Fig. 6C and D) (Supplementary data Figure S9B). Furthermore, the percent of total mRNAs associated with polyribosomes was similar between nucleolin-overexpressed and control Hey cell lines (fractions 6-10) (Supplementary data Figure S9C). Of the total CSF-1 mRNA in control Hey cells (GFP) and nucleolin-overexpressed Hey cells (NCL-GFP),

42.72±5.99% and 40.57±4.15%, respectively, were associated with polyribosomes (fractions 6-10) (Fig. 6C and D).

The long poly(A)_{n≥27} mRNAs associated with polyribosomes were obtained by LM-PAT-RT-qPCR analysis. There were significant differences in the long poly(A)_{n≥27} mRNAs associated with polyribosomes between two cell lines (*i.e.*, NCL-GFP and GFP Hey cell lines) ($P < 0.001$ in Fig. 6E and F). The products of LM-PAT-RT-qPCR fractionated by agarose gel electrophoresis show that CSF-1 mRNA is significantly deadenylated in Hey (NCL-GFP) cell line compared to Hey (GFP) cell line (Fig. 6G and H).

In nucleolin-overexpressed Hey cells (NCL-GFP), 17.26±1.64% of CSF-1 mRNA was long poly(A)_{n≥27} mRNAs associated with polyribosomes (fractions 6-10) (Fig. 6I). However, in control Hey cells (GFP), 37.96±4.65% of CSF-1 mRNA was long poly(A)_{n≥27} mRNAs associated with polyribosomes (fractions 6-10) ($P = 0.0019$ between two cell lines) (Fig. 6J). Since nucleolin mediates miRNA-directed deadenylation [17], the percent of deadenylated mRNAs (*i.e.*, short poly(A)_{n≤26} mRNAs) associated with polyribosomes was calculated. In nucleolin-overexpressed Hey cells (NCL-GFP), 23.31% of deadenylated CSF-1 mRNA was associated with polyribosomes, (Fig. 6I), while 4.76% of deadenylated CSF-1 mRNA (GFP) was associated with polyribosomes in control Hey cells (GFP) (Fig. 6J).

To verify the poly(A) tail length of CSF-1 mRNA, LM-PAT products were sequenced. In nucleolin overexpressed Hey cell line, the average poly(A) length was 18.6±1.1. In contrast, in control Hey cell line, the average poly(A) length was 109.6±3.0.

This data indicates that nucleolin increases the level of CSF-1 deadenylated mRNAs (*i.e.*, short poly(A)_{n≤26} mRNAs) associated with polyribosomes without significantly affecting total mRNA levels.

Since nucleolin down-regulation did not affect the poly(A)_n tail length of CSF-1 mRNA [17], LM-PAT analysis in nucleolin down-regulated cells was not performed.

4. Discussion

Cytoplasmic form of nucleolin

Even though both cyto-NCL and nu-NCL are equally abundant in immunoblot analysis (Fig. 1B), the ability to observe cyto-NCL under the immunofluorescence microscope is dependent on the region of nucleolin used to raise the antibody (Fig. 1A). We reason this could be due to masking protein(s) which block the individual antibody's access to nucleolin under immunofluorescence conditions. A masking protein(s) may bind to the nucleolin domain and block antibody accessibility. For example, the masking of nuclear export signals (NES) in both BRCA2 and its cargo protein, RAD51, localizes them in the nucleus [31]. Cancer-associated point mutations in NES in BRCA2 hinder the binding of the DSS1 masking protein, rendering BRCA2 cytoplasmic. The idea of masking protein(s), which bind the C- and/or N-terminal nucleolin and block antibody accessibility, is supported by the different results between immunoblot (Fig. 1B) and IF microscopy (Fig. 1E-G, Supplementary Data Figure S3-5). The N-terminal raised antibody detects both cyto-NCL and nu-NCL in denaturing immunoblot (*i.e.*, SDS-PAGE), but only detects the cyto-NCL in IF microscopy, *i.e.*, cells are formaldehyde-fixed and nu-NCL is not accessible by N-terminal raised antibody in IF microscopy. Similarly, the C-terminal raised antibody detects NCL in the nucleus only, primarily in the nucleolus. The cyto-NCL and nu-NCL are the same protein with the same MW and pI (Fig. 1B and D).

With this data, we propose that the 'C-terminal' end of the cyto-NCL is masked by unknown protein(s), while the 'N-terminal' end of the nu-NCL is also masked by unknown protein(s). Putative NES and NLS in the N- and C-terminals of nucleolin are likely masked by unknown proteins, resulting in the differential subcellular localizations of nucleolin; *i.e.*, cyto-NCL and nu-NCL. We hypothesize that the protein(s) which mask the C-terminal of the cyto-

NCL may be involved in nucleolin-mediated, miRNA-directed deadenylation and also translation. The masking proteins for cyto-NCL may include eIF4G, as detected by IP analysis (Fig. 4) and supported with our finding that cytoplasmic nucleolin co-localizes with eIF4G (Fig. 5).

The distinct foci formation of the cyto-NCL (Fig. 1), which is co-localized with the eIF4G (Fig. 5), in ovarian and breast cancer cells implicates their role in the increased level of CSF-1 expression in those cancer cell lines [32].

Balancing translation and mRNA decay by nucleolin

Nucleolin has a stretch of acidic residues in the N-terminal, four RRM domains in the mid region, and nine RGG repeats in the C-terminal. Each region has a distinct role in nucleolin function. The acidic region is involved in protein-protein interaction, the RRM domain for RNA-binding, and RGG repeats for RNA-binding [33] or protein-protein interaction [34]. The RNA-binding domain (RBD3 and RBD4) and the C-terminal RGG repeats are also known to accelerate the annealing of mismatched nucleic acids [35]. Co-IP data indicates that only the N-terminal antigen raised antibody co-immunoprecipitates nucleolin with eIF4G (Fig. 4), indicating that the middle domain and C-terminal of nucleolin are likely occupied by unknown protein(s) in the cytoplasm. Furthermore, N-terminal antigen raised antibody cannot immunoprecipitate nucleolin in RNases-untreated material, suggesting a tight mRNP complex formation. We propose that binding of nucleolin to the hI structure in the 5'UTR and eIF4G facilitates the eIF4F complex formation and enhances translation (Fig. 7).

We previously reported that nucleolin mediates miRNA-directed deadenylation by facilitating the miRISC formation in the 3'UTR of CSF-1 mRNA [17]. By doing this, nucleolin

may mediate the tight control of translation of CSF-1 mRNA containing short poly (A)_{n≤26} tails. This model is supported by the observation of the increased CSF-1 mRNA with short poly (A)_{n≤26} tails associated with polyribosomes in nucleolin overexpressed cells (Fig. 6).

The presence of a hairpin loop structure (hl) at the extreme end of the 5'UTR of CSF-1 mRNA enhances translation (Fig. 3). In contrast, miRNA target, G-quadruplex, and AREs in the 3'UTR of CSF-1 mRNA are involved in nucleolin-mediated miRNA-directed deadenylation. Our data indicates that binding of nucleolin to the hl in the 5'UTR, along with its association and co-localization with eIF4G, is primarily involved in CSF-1 mRNA translational regulation. Binding of nucleolin to the hl structure in the 5'UTR may ease the entry of eIF4F complex for translation initiation. Furthermore, the low dG (-7.7 kcal/mol) of hl in the 5'UTR of CSF-1 mRNA may be easily unwound by eIF4A during translation initiation. In contrast, *cis*-acting elements (*i.e.*, miRNA target, G-quadruplex, and AREs) in the 3'UTR of CSF-1 mRNA are mainly involved in mRNA deadenylation [17]. This suggests that the 5' hl structure, more so than the 3'UTR elements, is the primary regulator by which nucleolin promotes translation.

5. Conclusion

Our study indicates that multiple factors are involved in the dynamic formation of a CSF-1 mRNP structurome for translational regulation. These factors include *cis*-acting elements and RNA binding proteins. The interaction of nucleolin with *cis*-acting elements may balance mRNA decay and translation to maintain the proper level of CSF-1 protein in the cell. Disruption of this balance can lead to the abnormal expression of CSF-1 protein. Furthermore, the formation of distinct cytoplasmic foci of nucleolin in breast and ovarian cancer cells (Fig. 1) implicates the translational enhancing role of nucleolin in these cancer cells.

Competing interests

The authors declare that they have no competing interests.

Authors' Contribution

H.H.W. conceived the study, designed and performed the experiments, interpreted the results and wrote the manuscript. S.L. and S.G. performed the experiments. S.K.C. designed the experiments, interpreted the results, and wrote the manuscript.

Acknowledgments

We thank Stephen Greco, Michael Diri, and Arpan Patel for technical help in preparing plasmids and cell lines. This work was supported by the Women's Cancers of the University of Arizona Cancer Center, the University of Arizona Cancer Center Support Grant P30 CA023074, and the Bobbi Olson Endowment Fund (to SKC).

Appendix A. Supplementary data

Supplementary data to this article can be found on line.

References

- 1) L. Barret, S. Fletcher, S.D. Wilton, Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. *Cell Mol. Life Sci.* 69 (2012) 3613-3634.
- 2) F. Gebauer, M.W. Hentze, Molecular mechanisms of translational control. *Nat. Rev. Mol. Cell. Biol.* 5 (2004) 827-835.
- 3) S.E. Dmitriev, D.E. Andreev, Z.V. Adyanova, I.M. Terenin, I.N. Shatsky, Efficient Cap-Dependent Translation of Mammalian mRNAs with Long and Highly Structured 5'-Untranslated Regions in Vitro and in Vivo. *Mol. Biol.* 43 (2009) 108-113.
- 4) P. Koukouraki, E. Doxakis, Constitutive translation of human α -synuclein is mediated by the 5'-untranslated region. *Open Biol.* 6 (2016) 160022.
- 5) J.R. Babendure, J.L. Babendure, J.H. Ding, R.Y. Tsien, Control of mammalian translation by mRNA structure near caps. *RNA* 12 (2006) 851-861.
- 6) B.M. Pickering, A.E. Willis, The implications of structured 5' untranslated regions on translation and disease. *Sem Cell Dev. Biol.* 16 (2005) 39-47.
- 7) C.Y.A. Chen, A.B. Shyu, Mechanisms of deadenylation-dependent decay. *Wiley Interdiscip. Rev.* 2 (2011) 167-183.
- 8) C.Y.A. Chen, D. Zheng, Z. Xia, A.B. Shyu, Ago-TNRC6 complex triggers microRNA-mediated mRNA decay by promoting biphasic deadenylation followed by decapping. *Nat. Struct. Mol. Biol.* 16 (2010) 1160–1166.
- 9) C.P. Wigington, K.R. Williams, M.P. Meers, G.J. Bassell, A.H. Corbett, Poly(A) RNA binding proteins and polyadenosine RNA: New members and novel functions. *Wiley Interdiscip. Rev. RNA* 5 (2014) 601–622.

- 10) T.E. Dever, R. Green, The elongation, termination, and recycling phases of translation in eukaryotes. *Cold Spring Harb. Perspect. Biol.* 4:a013706 (2012).
- 11) S.Z. Jr. Tarun, A.B. Sachs, Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO J.* 15 (1996) 7168-7177.
- 12) V. Pelechano, W. Wei, L.M. Steinmetz, Widespread co-translational RNA decay reveals ribosome dynamics. *Cell* 161 (2015) 1400-1412.
- 13) A. Yamashita, T.C. Chang, Y. Yamashita, W. Zhu, Z. Zhong, C.Y. Chen, A.B. Shyu, Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nat. Struct. Mol. Biol.* 12 (2005) 1054-1063.
- 14) C.Y. Chen, A.B. Shyu, Deadenylation and P-bodies. *Adv Exp Med Biol* 768 (2013)183-195.
- 15) N. Xu, C.Y. Chen, A.B. Shyu, Modulation of the fate of cytoplasmic mRNA by AU-rich elements: key sequence features controlling mRNA deadenylation and decay. *Mol. Cell. Biol.* 17 (1997) 4611-4621.
- 16) D. Zheng, N. Ezzeddine, C.Y. Chen, W. Zhu, X. He, A.B. Shyu, Deadenylation is prerequisite for P-body formation and mRNA decay in mammalian cells. *J. Cell Biol.* 182 (2008) 89-101.
- 17) H.H. Woo, T. Baker, C. Laszlo, S.K. Chambers, Nucleolin mediates microRNA-directed CSF-1 mRNA deadenylation but increases translation of CSF-1 mRNA. *Mol. Cell. Proteomics* 12 (2013) 1661-1677.
- 18) J. Chen, K. Guo, M. Kastan, Interactions of nucleolin and ribosomal protein L26 (RPL26) in translational control of human p53 mRNA. *J. Biol. Chem.* 287 (2012) 16467-16476.

- 19) T.K. Sengupta, S. Bandyopadhyay, D.J. Fernandes, E.K. Spicer, Identification of nucleolin as an AU-rich element binding protein involved in bcl-2 mRNA stabilization. J.B.C. 279 (2004) 10855-10863.
- 20) D. Finger, L. Trantirek, C. Johansson, J. Feigon, Solution structures of stem-loop RNAs that bind to the two N-terminal RNA-binding domains of nucleolin. Nucleic Acids Res. 39 (2003) 8513–8530.
- 21) P. Bouvet, F.H. Allain, L.D. Finger, T. Dieckmann, J. Feigon, Recognition of pre-formed and flexible elements of an RNA stem-loop by nucleolin. J. Mol. Biol. 309 (2001) 763-775.
- 22) S.K. Chambers, Role of CSF-1 in progression of epithelial ovarian cancer. Future Oncol. 5 (2009) 1429-1440.
- 23) T.K. Sengupta, S. Bandyopadhyay, D.J. Fernandes, Spicer EK, Identification of nucleolin as an AU-rich element binding protein involved in bcl-2 mRNA stabilization. J.B.C. 279 (2004) 10855-10863.
- 24) H.H. Woo, C. Laszlo, S. Greco, S.K. Chambers, Regulation of colony stimulating factor-1 expression and ovarian cancer cell behavior in vitro by miR-128 and miR-152. Mol. Cancer 11 (2012) 58.
- 25) E.L. Murray, D.R. Schoenberg, Assays for determining poly(A) tail length and the polarity of mRNA decay in mammalian cells. Methods Enzymol. 448 (2008) 483-504.
- 26) H.H. Woo, X. Yi, T. Lamb, I. Menzl, T. Baker, D.J. Shapiro, S.K. Chambers, Posttranscriptional suppression of proto-oncogene *c-fms* expression by vigilin in breast cancer. Mol. Cell. Biol. 31(2011) 215-225.

- 27) N. Kedersha, P. Anderson, Mammalian stress granules and processing bodies. *Methods in Enzymology* 431 (2007) 61-81.
- 28) C.M. Berger, X. Gaume, P. Bouvet, The roles of nucleolin subcellular localization in cancer. *Biochimie* 113 (2015) 78-85.
- 29) B.W. Baer, R. Kornber, Repeating structure of cytoplasmic poly(A)-ribunucleoprotein. *P.N.A.S.* 77 (1980) 1890-1892.
- 30) J.E. Park, H. Yi, Y. Kim, H. Chang, V.N. Kim, Regulation of poly(A) tail and translation during the somatic cell cycle. *Mol. Cell* 62 (2016) 462-471.
- 31) A.D. Jeyasekharan, Y. Liu, H. Hattori, V. Pisupati, A.B. Jonsdottir, E. Rajendra, M. Lee, E. Sundaramoorthy, S. Schlachter, C.F. Kaminski, Y. Ofir-Rosenfeld, K. Sato, J. Savill, N. Ayoub, A.R. Venkitaraman, A cancer-associated BRCA2 mutation reveals masked nuclear export signals controlling localization. *Nat. Struct. Mol. Biol.* 20 (2013) 1191-1198.
- 32) S.K. Chambers, B.M. Kacinski, C.M. Ivins, M.L. Carcagiu, Overexpression of epithelial macrophage colony-stimulating factor (CSF-1) and CSF-1 receptor: A poor prognostic factor in epithelial ovarian cancer, contrasted with a protective effect of stromal CSF-1. *Clin. Cancer Res.* 3 (1997) 999-1007.
- 33) S. Calabretta, S. Richard, Emerging roles of disordered sequences in RNA-binding proteins. *Trends Biochem. Sci.* 1180 (2015) 1-11.
- 34) P. Rajyaguru, M. She, R. Parker, Scd6 targets eIF4G to repress translation: RGG motif proteins as a class of eIF4G-binding proteins. *Mol. Cell* 45(2012) 244-254.
- 35) L.A. Hanakahi, Z. Bu, N. Maizels, The C-terminal domain of nucleolin accelerates nucleic acid annealing. *Biochem.* 39 (2000) 15493-15499.

- 36) C. Wilusz, M. Wormington, S. Peltz, The cap-to-tail guide to mRNA turnover. *Nat. Rev. Mol. Cell Biol.* 2 (2001) 237-246.
- 37) A.W.B. Craig, A. Haghighat, A.T. Yu, N. Sonenberg, Interaction of polyadenylate-binding protein with the eIF4G homologue PAIP enhances translation. *Nature* 392 (1998) 520-523.
- 38) K. Abdelmohsen, K. Tominaga, E.K. Lee, S. Srikantan, M.J. Kang, M.M. Kim, R. Selimyan, J.L. Martindale, X. Yang, F. Carrier, M. Zhan, K.G. Becker, M. Gorospe, Enhanced translation by Nucleolin via G-rich elements in coding and non-coding regions of target mRNAs. *Nucleic Acids Res.* 39 (2011) 8513-8530.

Figure legends

Fig. 1. Cytoplasmic form of nucleolin.

- (A) The N-terminal, middle domain, and C-terminal immunogens used to raise nucleolin antibodies.
- (B) Nucleolin immunoblot of cytoplasmic and nuclear fractions by antibodies raised against the N-terminal, middle domain, and C-terminal immunogens. All antibodies detect a single band with a MW $\cong$ 110 kDa in SDS-PAGE.
- (C) Actin, HDAC1, and Pan3 immunoblots show purity of cytoplasmic and nuclear lysates of Hey ovarian cancer cells.
- (D) Nucleolin immunoblot of cytoplasmic and nuclear fractions detects a single band with a pI $\cong$ 4.6 in IEF.
- (E) Subcellular localization of nucleolin by indirect immunofluorescence microscopy in MCF10A non-cancerous epithelial breast cells. The N-terminal raised antibody does not detect nucleolin in either the cytoplasm or nucleus. The middle domain raised antibody detects nucleolin both in the cytoplasm and nucleus. The C-terminal raised antibody detects nucleolin in the nucleus only.
- (F) Subcellular localization of nucleolin in MDA-MB-231 breast cancer cells. The N-terminal raised antibody detects nucleolin in the cytoplasm only. The middle domain raised antibody detects nucleolin both in the cytoplasm and nucleus. The C-terminal raised antibody detects nucleolin only in the nucleus, primarily in the nucleolus. Nucleolin shRNA decreased nucleolin signal in the cytoplasm detected by the N-terminal raised antibody.

(G) The cytoplasmic form of nucleolin in MDA-MB-231 cells detected by the N-terminal and middle domain raised antibodies is co-localized. The bar is 10 μ m.

Fig. 2. The extreme 5' end sequence in the 5'UTR of CSF-1 mRNA encoding the hl structure is responsible for translational regulation.

(A) Luciferase constructs with CSF-1 mRNA 5' and 3'UTRs. The number on top indicates the *cis*-acting elements' locations in the CSF-1 mRNA sequence. All constructs are under the control of a tetracycline-inducible promoter (Tet-ON-3G). All constructs are chromosomally-integrated in SKOV3 and Hey cells for stable expression. Solid box depicts wild type sequence. Open box depicts mutation.

(B) Significant increase of luciferase activity in the Wt=Mut(Tripl) construct compared to the other constructs tested in SKOV3 cells. The luciferase activity increases dramatically in the triple mutant {Wt= Mut(Tripl)}. However, deletion of the tip in the 5'UTR in the triple mutant revokes the effect of the triple mutant by decreasing luciferase activity dramatically. Y axis indicates fold change of luciferase activity.

(C) Similar effect is observed in Hey cells.

(D) Ratio of luciferase activity divided by the induction time in SKOV3 cells. Y axis indicates ratio (L/t).

(E) Ratio of luciferase activity divided by the induction time in Hey cells.

(F) Compared to the mRNA level of the Wt=Wt construct, mRNA levels of the Wt=Mut(Tripl) and D(21)=Mut(Tripl) constructs are ~4.5-fold higher in SKOV3 cells after 4 hr DOX removed in SKOV3 cells. Compared to the mRNA level of the Wt=Wt construct, mRNA level of the D(21)=Wt construct is ~13-fold lower after 4 hr DOX removed in SKOV3 cells.

(G) Compared to the mRNA level of the Wt=Wt construct, mRNA levels of the Wt=Mut(Tripl) and D(21)=Mut(Tripl) constructs are ~5-fold higher in Hey cells after 4 hr DOX removed in Hey cells. Compared to the mRNA level of the Wt=Wt construct, mRNA level of the D(21)=Wt construct is ~2.5-fold lower after 4 hr DOX removed in Hey cells. Mean $\pm$ S.D. is presented. N = 10.

Fig. 3. Nucleolin interacts with the hl structure in the 5'UTR as well as G-quadruplex in the 3'UTR of CSF-1 mRNA.

- (A) RNA electrophoretic mobility shift assay shows an association of nucleolin (tagged with myc-DDK) with the hl structure from the 5'UTR as well as with the G-quadruplex from the 3'UTR of CSF-1 mRNA. G-quadruplex is used as a nucleolin binding control [17]. Nucleolin does not associate with either the linear 5'hl sequence or 5'UTR without hl sequence.
- (B) Luciferase protein level is increased (4~5-fold) in wild type construct (Wt=Wt) in response to the nucleolin overexpression in SKOV3 cells. In contrast, luciferase protein level is not increased in the hl deletion construct {D(21)=Wt} in response to the nucleolin overexpression.
- (C) Luciferase mRNA stability is decreased in D(21)=Wt (red solid line) compared to Wt=Wt (black solid line) in SKOV3 cells. However, there was no significant effect of nucleolin overexpression on luciferase mRNA stability (broken lines). The cDNA was synthesized by random primer (pdN₆). This suggests an effect of nucleolin on luciferase mRNA translation.

Fig. 4. Nucleolin forms mRNP complex with eIF4G and CSF-1 mRNA in Hey ovarian cancer cells.

- (A) N-terminal raised antibody (N-NCL) immunoprecipitates nucleolin only in RNases-treated sample. In the same sample, nucleolin co-immunoprecipitates with eIF4G.
- (B) Middle domain raised antibody (M-NCL) immunoprecipitates nucleolin in all treatments. However, nucleolin did not co-immunoprecipitate with eIF4G in the same samples.
- (C) C-terminal raised antibody (C-NCL) immunoprecipitates nucleolin. However, nucleolin did not co-immunoprecipitate with eIF4G in the same samples.
- (D) eIF4G antibody immunoprecipitates eIF4G in all treatments. However, eIF4G co-immunoprecipitates with nucleolin only in RNases-treated sample.
- (E) RT-qPCR measurements of CSF-1 mRNA in nucleolin or eIF4G immunoprecipitates of Hey cell lysates show direct interaction between either nucleolin or eIF4G and CSF-1 mRNA. After IP of RNA-protein complexes from Hey cell lysates, RNA was isolated and used in RT reactions and amplified by real time PCR. The graph shows relative mRNA levels in either nucleolin or eIF4G IP compared with control IgG IP conditions. The mean $\pm$ SD of CSF-1 mRNA normalized for GAPDH mRNA is depicted ($n = 3$). GAPDH mRNA was set to equal 1.

Fig. 5. Cytoplasmic nucleolin co-localizes with the eIF4G in MDA-MB-231 breast cancer cells. The cytoplasmic nucleolin (green) detected by the N-terminal raised antibody co-localized with eIF4G (red). The bar is 10 μ m.

Fig. 6. Increased level of the CSF-1 mRNA with short poly (A) tail is associated with polyribosomes in nucleolin overexpressed Hey ovarian cancer cell cytoplasm.

- (A) Ribosome profiles of Hey (NCL-GFP) and Hey (GFP) cytoplasmic lysates.
- (B) Distribution of RBPs in ribosome profile in Hey (NCL-GFP) cells.

- (C) Ribosome profile of CSF-1 total mRNA from Hey (NCL-GFP) cytoplasmic lysates.
- (D) Ribosome profile of CSF-1 total mRNA from Hey (GFP) cytoplasmic lysates.
- (E) Ribosome profile of CSF-1 long poly(A)_{n≥27} mRNA from Hey (NCL-GFP) cytoplasmic lysates.
- (F) Ribosome profile of CSF-1 long poly(A)_{n≥27} mRNA from Hey (GFP) cytoplasmic lysates.
- (G) DNA blot of LM-PAT-RT-qPCR products from Hey (NCL-GFP) fractionated in agarose gel electrophoresis. Most CSF-1 mRNAs have short poly(A)_{n≤26} tails.
- (H) DNA blot of LM-PAT-RT-qPCR products from Hey (GFP) fractionated in agarose gel electrophoresis. Most CSF-1 mRNAs have long poly (A)_{n≥27} tails.
- (I) In Hey (NCL-GFP) cells, 23.31% of short poly(A)_{n≤26} CSF-1 mRNA and 17.26% of long poly(A)_{n≥27} CSF-1 mRNA were associated with polyribosomes (fractions 6-10)
- (J) In Hey (Empty) cells, 4.76% of short poly(A)_{n≤26} CSF-1 mRNA and 37.96% of long poly(A)_{n≥27} CSF-1 mRNA were associated with polyribosomes (fractions 6-10). Mean ± S.D. is presented. N = 6.

Fig. 7. Formation of a nucleolin-mediated mRNP in translational regulation.

Binding of nucleolin to the hI structure in the 5'UTR attracts eIF4F complex formation and enhances translation. Binding of nucleolin to the 3'UTR facilitates miRISC formation resulting in miRNA-directed deadenylation, but blocks further mRNA decay.

Fig. 1.

F

MDA-MB-231

G

Fig. 2. A

Fig. 3.

Fig. 4.

Fig. 5.

Fig. 6. A

B

C

D

Fig. 7.

*Binding of nucleolin to hairpin loop structure eases entry of the eIF4F complex

*Binding of nucleolin to the G-quadruplex/ AREs facilitates miRISC formation, but blocks further mRNA decay

Supplementary Data

Expression of the cytoplasmic nucleolin for post-transcriptional regulation of CSF-1 mRNA in ovarian and breast cancer cells

Ho-Hyung Woo, Sang C. Lee, Steven J. Gibson, Setsuko K. Chambers

From the University of Arizona Cancer Center, Tucson, AZ 85724, USA

Figure S1. Nucleolin immunoblot of the cytoplasmic fractions of Hey ovarian cancer cells transfected either by nucleolin shRNA (lanes 1,3,5) or scrambled shRNA (lanes 2,4,6). N-terminal, middle domain, of C-terminal antigen raised antibodies were used.

Figure S2. Nucleolin immunoblot of cytoplasmic and nuclear fractions. Tubulin, HDAC1, and actin immunoblots show the purity of cytoplasmic and nuclear lysates.

Figure S3. Subcellular localization of nucleolin by indirect immunofluorescence microscopy in BT20 breast cancer cells. The N-terminal raised antibody detects nucleolin in the cytoplasm. The middle domain raised antibody detects both in the cytoplasm and the nucleus. The C-terminal raised antibody detects nucleolin in the nucleus. The bar is 10 μm .

Figure S4. Subcellular localization of nucleolin by indirect immunofluorescence microscopy in Hey ovarian cancer cells. The N-terminal raised antibody detects nucleolin in the cytoplasm. The middle domain raised antibody detects nucleolin in the nucleus. The C-terminal raised antibody detects nucleolin in the nucleus. The bar is 10 μm .

Figure S5. Subcellular localization of nucleolin by indirect immunofluorescence microscopy in SKOV3 ovarian cancer cells. The N-terminal raised antibody detects nucleolin in the cytoplasm. The middle domain raised antibody detects nucleolin in the nucleus. The C-terminal raised antibody detects nucleolin in the nucleus. The bar is 10 μm .

Figure S6. Hairpin loop structure at the extreme 5' end of the 5'UTR in CSF-1 mRNA is generated by RNAfold program. 2-21nt is involved in hairpin loop structure formation. $\Delta G = -7.7$ kcal/mol.

Figure S7. Determination of % of deadenylated mRNAs associated with polyribosomes in total mRNAs by the sucrose density gradient fractionation and LM-PAT-RT-qPCR.

Figure S8. The distribution of GAPDH total mRNA in ribosome profile in A) Hey (NCL-GFP), and B) Hey (GFP) cells. The distribution of GAPDH poly (A)_{n>27} mRNA in C) Hey (NCL-GFP)

and D) Hey (GFP) cells. E and F) DNA blot of LM-PAT-RT-qPCR products fractionated in agarose gel electrophoresis. The GAPDH deadenylated mRNA (poly (A)_{n<26}) associated with polyribosomes (fraction 6-10) is G) 4.74% in Hey (NCL-GFP) and H) 5.66% in Hey (GFP) cells.

A

B

C

Figure S9. Immunoblot of A) CSF-1 and actin in Hey (GFP) and Hey (NCL-GFP) cytoplasmic extracts. CSF-1 protein level is increased in Hey (NCL-GFP) cells. Serial increase of total protein is loaded for comparison. B) There are no changes in relative CSF-1 mRNA levels between Hey (GFP) and Hey (NCL-GFP) cells. C) There are no changes in relative CSF-1 total mRNA levels associated with heavy polyribosomes (fractions 6-10) between Hey (GFP) and Hey (NCL-GFP) cells.