Nucleolar stress and impaired stress granule formation contribute to C9orf72 RAN translation-induced cytotoxicity

Zhouteng Tao¹, Hongfeng Wang¹, Qin Xia¹, Ke Li¹, Kai Li², Xiaogang Jiang², Guoqiang Xu³, Guanghui Wang¹,⁵,* and Zheng Ying¹,⁴,*

¹Laboratory of Molecular Neuropathology, Jiangsu Key Laboratory of Translational Research and Therapy for Neuro-Psycho-Diseases and College of Pharmaceutical Sciences, ²Department of Pharmacology, College of Pharmaceutical Sciences, ³Laboratory of Chemical Biology, Jiangsu Key Laboratory of Translational Research and Therapy for Neuro-Psycho-Diseases and College of Pharmaceutical Sciences, ⁴Jiangsu Key Laboratory of Preventive and Translational Medicine for Geriatric Diseases, College of Pharmaceutical Sciences, Soochow University, Suzhou, Jiangsu 215021, China and ⁵Key Laboratory of Brain Function and Disease, School of Life Sciences, University of Science and Technology of China, Chinese Academy of Sciences, Hefei, Anhui 230027, China

*To whom correspondence should be addressed at: Laboratory of Molecular Neuropathology, Jiangsu Key Laboratory of Translational Research and Therapy for Neuro-Psycho-Diseases and College of Pharmaceutical Sciences, Soochow University, Suzhou, Jiangsu 215021, China. Tel: +86 51265884845; Fax: +86 51265884845; Email: zheng.ying@suda.edu.cn (Z.Y.) or wanggh@suda.edu.cn (G.W.)

Abstract

Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) are the two common neurodegenerative diseases that have been associated with the GGGGCC-GGCCCC repeat RNA expansion in a noncoding region of C9orf72. It has been previously reported that unconventional repeat-associated non-ATG (RAN) translation of GGGGCC-GGCCCC repeats produces five types of dipeptide-repeat proteins (referred to as RAN proteins): poly-glycine-alanine (GA), poly-glycine-proline (GP), poly-glycine-arginine (GR), poly-proline-arginine (PR) and poly-proline-alanine (PA). Although protein aggregates of RAN proteins have been found in patients, it is unclear whether RAN protein aggregation induces neurotoxicity. In the present study, we aimed to understand the biological properties of all five types of RAN proteins. Surprisingly, our results showed that none of these RAN proteins was aggregate-prone in our cellular model and that the turnover of these RAN proteins was not affected by the ubiquitin–proteasome system or autophagy. Moreover, poly-GR and poly-PR, but not poly-GA, poly-GP or poly-PA, localized to the nucleolus and induced the translocation of the key nucleolar component nucleophosmin, leading to nucleolar stress and cell death. This poly-GR- and poly-PR-mediated defect in nucleolar function was associated with the suppression of ribosomal RNA synthesis and the impairment of stress granule formation. Taken together, the results of the present study suggest a simple model of the molecular mechanisms underlying RAN translation-mediated cytotoxicity in C9orf72-linked ALS/FTD in which nucleolar stress, but not protein aggregation, is the primary contributor to C9orf72-linked neurodegeneration.
Introduction

Recent studies showed that an abnormal expansion of a repeated hexanucleotide GGGGCC sequence in a noncoding region of the C9orf72 gene represents the most common cause of both amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD), two neurodegenerative disorders that present significant clinical and pathological overlap (1–3). Subsequent studies provided further evidence that C9orf72 hexanucleotide repeat expansion may also play roles in Alzheimer’s disease and Huntington’s disease (4,5), suggesting both the importance of C9orf72 and a common molecular pathway in various neurodegenerative disorders. Typically, normal individuals carry 2–23 GGGGCC repeats, whereas C9orf72-linked ALS/FTD (referred to as C9ALS/FTD) patients carry an expansion with dozens to thousands of repeats (1,2,6). However, the relationship between the number of repeats and disease onset or progression remains unclear.

Elucidating how the bidirectional transcription of expanded GGGGCC-GGCCGCC repeats leads to neurodegeneration and identifying effective therapeutic treatments for C9ALS/FTD have become the two major goals within the field of C9ALS/FTD research. To date, the following three major hypotheses for the mechanisms underlying this disease have been posited. (1) A protein loss-of-function mechanism may be caused by C9orf72 haploinsufficiency. Although little is known about C9orf72-coding protein function, an interesting recent study showed that the C9orf72 protein may function in endosomal trafficking and autophagy (7). Hexanucleotide repeat expansion might impair the expression of the coding product of C9orf72, and it has been widely demonstrated that the C9orf72 mRNA levels are reduced in patients carrying a C9orf72 expansion (1–3), suggesting that partial loss of C9orf72 transcript expression may be associated with disease pathogenesis. In support of this loss-of-function mechanism, studies using animal models, such as the vertebrate zebrafish and the invertebrate Caenorhabditis elegans, showed that C9orf72 depletion caused deficient locomotion and motor neuron degeneration (8,9). However, families carrying a C9orf72 expansion exhibit a dominant inheritance pattern, which suggests that loss of C9orf72 function may not be a key mechanism of the disease. (2) An RNA gain-of-function mechanism. To date, most efforts to understand the mechanism underlying C9ALS/FTD have exclusively focused on the contribution of RNA-mediated toxicity toward manifestation of the disease. The formation of RNA foci containing the repeat expansion is a key characteristic of many repeat expansion diseases associated with RNA-mediated toxicity (10). Both sense and antisense RNA foci of C9orf72 hexanucleotide repeats have been found in C9ALS/FTD patient cells and tissues (11–16), and sequestration and dysregulation of RNA-binding proteins by hairpin and G-quadruplex structures consisting of GGGGCC repeats impair RNA processing and contribute to disease pathogenesis (12,15,17–20). However, it remains unclear whether these RNA-binding proteins associate with endogenous GGGGCC repeats or RNA foci in vivo. (3) A protein gain-of-function mechanism. Repeat-associated non-ATG (RAN) translation was initially reported to be associated with spinocerebellar ataxia type 8 (SCA8) and myotonic dystrophy type 1 (DM1), both of which involve a CAG repeat expansion (21). Subsequently, RAN translation products of expanded GGGGCC and CGG repeats were identified in C9ALS/FTD (22,23) and fragile X-associated tremor ataxia syndrome, respectively (24). Shortly thereafter, several studies further demonstrated that the antisense GGCCCC repeat is also RAN translated (13,16,25). Depending on the reading frame of the bidirectional transcript, GGGGCC-GGCCGCC repeats encode poly-glycine-alanine (GA), poly-glycine-proline (GP), poly-glycine-arginine (GR), poly-proline-arginine (PR) and poly-proline-alanine (PA) dipeptide repeat-containing proteins, which are referred to as RAN proteins. Although the cytoplasmic and nuclear aggregation of RAN proteins in patient brains suggests that these proteins may be pathogenic (13,16,22,23,25–27), direct evidence for the role of RAN proteins in disease pathogenesis largely remains lacking. Recently and intriguingly, poly-GR and poly-PR were identified to be cytotoxic (28), most likely by interfering with RNA biogenesis in nucleoli (29), suggesting that RAN proteins may play important roles in disease pathogenesis.

Here, we aimed to comprehensively examine the biological properties of these five types of RAN proteins to determine whether and the mechanism by which they contribute to cytotoxicity. Surprisingly and unexpectedly, our results showed that none of the RAN proteins was aggregate-prone and that poly-GR and poly-PR, but not poly-GA, poly-GP or poly-PA, localized to the nucleolus and induced cell death in association with nucleophosmin (B23) translocation from the nucleolus to the nucleoplasm, impaired ribosome RNA biogenesis and stress granule formation. Taken together, the two elegant studies of Kwon et al. and Miziołinska et al. and our results highlight nucleolar stress as a fundamental biological determinant of neurodegeneration and reveal the molecular mechanism underlying RAN translation-mediated neurotoxicity in ALS/FTD.

Results

Cell death is differentially induced by poly-GR and poly-PR, but not poly-GA, poly-GP or poly-PA, RAN proteins or GGGGCC RNA repeats

To determine whether RAN proteins directly induce cytotoxicity, it was necessary to dissociate the potential toxicity of hexanucleotide RNA repeats from that of RAN protein repeats. Thus, we generated constructs encoding RAN proteins that contained non-GGGGCC repeats, such as GGAGCA repeats that encoded poly-GR and GGAAGA repeats that encoded poly-GA, which were referred to as ‘synthetic’ constructs. The synthetic constructs were predicted to lack the ability to produce RNA foci and other RAN proteins because they could not form putative hairpin structures, in contrast to GGGGCC-GGCCGCC repeats (Supplementary Material, Fig. S1). We also generated ATG start codon-initiated GGGGCC and CCGGG repeats that encoded all five RAN proteins using various reading frames, which were referred to as ‘pathogenic’ constructs (Fig. 1A).

All of the above-mentioned GFP-tagged constructs, each of which contained 30 repeat sequences, and a GFP-tagged construct containing a repeated GGGGCC sequence lacking an ATG (No ATG) were transfected into cells. The constructs (both synthetic and pathogenic) that encoded arginine-containing poly-dipeptides (poly-GR or poly-PR) strikingly induced cell death in both neuronal (NSC-34) and non-neuronal (HEK293) cells (Fig. 1B and C). Given that the GGGGCC RNA repeats alone (No ATG) exhibited very weak toxicity following an extended transfection period compared with poly-GR and poly-PR (Fig. 1D and E), we reasoned that the five RAN proteins encoded by the pathogenic constructs containing ATG-preceded GGGGCC and CCGGG repeats should display toxicity following a standard transfection period. In support of this notion, the GGGGCC- and non-GGGGCC-encoded RAN proteins, such as poly-GA and poly-GR, exerted similar cytotoxic effects (Fig. 1B and C). These observations were further confirmed by the MTT assay (Supplementary Material, Fig. S2).
Figure 1. Selective cell death induced by poly-GR and poly-PR, but not poly-GA, poly-GP, poly-PA RAN proteins or GGGGCC RNA repeats alone. (A) Schematic diagram of the hexanucleotide repeat constructs used in this study. Essentially, these constructs contained 30 synthetic or pathogenic hexanucleotide repeats. For example, GAggagca indicates GA(ggggcc)30. We termed the constructs with 60 hexanucleotide repeats 'long'. For example, GAggagca-long indicates GA(ggggcc)60. (B and C) NSC-34 or HEK293 cells were transfected with the GFP tag alone or a GFP-tagged RAN construct as indicated. Thirty-six hours later, the cells were subjected to PI staining to count the dead cells. The quantitative data are indicated as the means ± SEM, *P < 0.05, **P < 0.01. (D) NSC-34 or HEK 293 cells were transfected with GFP or a GFP-tagged GGGGCC lacking an ATG (No ATG) for 36 or 72 h. Then, the cells were subjected to PI staining. The quantitative data are indicated as the means ± SEM, *P < 0.05. (E) NSC-34 cells were transfected with GFP, GFP-GRggagca, GFP-GRggaaga or GFP-PRccccgg. Cell death was analyzed via PI staining 72 h later. The quantitative data are indicated as the means ± SEM, **P < 0.01. (F) NSC-34 cells were transfected with GFP, GFP-GRggaaga or GFP-GRggaaga containing a TAG stop codon, and 36 h later, the cells were subjected to PI staining. The quantitative data are indicated as the means ± SEM, **P < 0.01. (G) HEK293 cells were transfected with FLAG, FLAG-GAggagca or FLAG-GAggagca-long (containing 60 repeats), FLAG-GRggaaga or FLAG-GRggaaga-long (containing 60 repeats). Thirty-six hours later, cell death was measured via PI staining, and the cell morphology was shown using PH microscopy. The quantitative data of cell death are indicated as the means ± SEM, *P < 0.05.
To further confirm that cytotoxicity was initiated by the RAN protein repeats but not the RNA repeats, we added a TAG stop codon 5’ to the GGAAGA sequence. This manipulation completely abolished the cytotoxicity induced by synthetic poly-GR (Fig. 1F), indicating that the synthetic poly-GR used in our study perfectly mimicked the toxic properties of the pathogenic poly-GR. To determine whether the cytotoxicity induced by arginine-containing RAN proteins is repeat length-dependent and to exclude the possibility that the artificial tag was responsible for the effect, we transfected cells with FLAG-tagged constructs encoding 60 repeats of synthetic poly-GA or poly-GR (referred to as ‘long’) and found that the 60-repeat poly-GR displayed higher cytotoxicity than the 30-repeat poly-GR. However, neither the 60-repeat nor the 30-repeat poly-GA had any effect on cytotoxicity (Fig. 1G). Similar observations were obtained using NSC-34 cells (data not shown), suggesting that longer repeat expansions in arginine-containing RAN proteins induce more cytotoxicity.

The arginine-enriched RAN proteins poly-GR and poly-PR form ‘aggregate’-like inclusion bodies in the nucleus

To determine the physiological role of these RAN proteins, we transfected HEK293 and neuronal cells with GFP- or FLAG-tagged RAN proteins and examined the cellular localization of those RAN proteins using fluorescence microscopy. The results showed that poly-GA, poly-GP and poly-PA were diffusively expressed throughout the cell (in both the nucleus and cytoplasm), whereas poly-GR and poly-PR, the two arginine-enriched RAN proteins, formed structures resembling protein aggregates in the nucleus, similar to a previous observation from polyglutamine (polyQ) inclusion bodies (Fig. 2A, B and E). Notably, pathogenic and synthetic poly-GA and poly-GR displayed similar subcellular localization (Fig. 2A, D and E), which was not affected by the tag or the repeat length (Supplementary Material, Fig. S3). In addition, we found that poly-GR could also diffusely distribute in the cytoplasm and partially diffuse in the nucleus (Fig. 2A and C), whereas poly-PR displayed an exclusively nuclear distribution (Fig. 2A, C and E).

None of the five RAN proteins or nuclear ‘aggregates’ formed by poly-GR and poly-PR are highly regulated by the ubiquitin–proteasome system (UPS) or autophagy. Because all RAN proteins have been found in aggregates in C9ALS/FTD brains and because the most remarkable characteristic of protein aggregates is that they are tightly associated with and regulated by the UPS or autophagy (30–32), we examined whether these ‘aggregate’-like inclusion bodies were formed by misfolded poly-dipeptides, similar to polyQ protein aggregates, and whether UPS and autophagy targeted these inclusion bodies for protein degradation. Therefore, we transfected cells with RAN proteins and treated the cells with the proteasomal inhibitor MG132 and the autophagy inhibitor Bafilomycin A1 (Baf A1). Surprisingly, no change in the number of nuclear ‘aggregates’ or in the protein levels of poly-GR or poly-PR was observed after this treatment (Fig. 3A and B), which was consistent with the finding that the nuclear ‘aggregates’ formed by poly-GR and poly-PR were not changed in Atg5 knockout (autophagy-deficient) mouse embryonic fibroblasts (MEFs) compared with Atg5 wild-type (autophagy-competent) MEFs (Fig. 3D). Additionally, the other three RAN proteins were not generally degraded via the UPS or autophagy (Fig. 3B) and were not aggregate-prone, as inhibiting the UPS and autophagy failed to induce their aggregation (Fig. 3A). We also found that the 60-repeat RAN proteins were not affected by the inhibition of the UPS and autophagy (data not shown). In contrast, the number of polyQ protein (Htt-60Q) aggregates was strikingly increased in MG132-treated cells (Fig. 3C) and in Atg5 knockout cells (data not shown). Taken together, our results suggest that neither the proteasomal nor the autophagy pathway strongly regulates the formation of nuclear inclusion bodies by the two arginine-containing RAN proteins or any other RAN proteins.

The nuclear inclusion bodies formed by poly-GR and poly-PR localize to the nucleoli but are not protein aggregates

As the nuclear inclusion bodies were not affected by the UPS or autophagy, the characteristics of these nuclear inclusions remained unclear. Next, we examined whether these inclusions are associated with the markers of protein aggregates, such as ubiquitin. We also examined VCP/p97 (VCP) and optineurin (OPTN) because these proteins have been well characterized to be associated with ALS/FTD and to be involved in protein quality control systems, such as the UPS and autophagy. As shown in Supplementary Material, Figure S4, none of these proteins colocalized with the nuclear inclusion bodies composed of poly-GRs (Supplementary Material, Fig. S4A–C). In contrast, the nuclear inclusions composed of full-length and truncated ataxin-3 containing an expanded polyQ sequence clearly colocalized with ubiquitin (Supplementary Material, Fig. S4D). These data suggest that the nuclear inclusion bodies formed by arginine-containing RAN proteins are not actual protein aggregates.

To better understand the properties of these inclusions, we evaluated the morphological characteristics of the nuclear inclusion bodies formed by poly-GR using phase contrast (PH) microscopy. We found that these inclusions localized to the nucleoli, which appeared dark under PH conditions (Fig. 4A). In contrast, nuclear inclusion bodies composed of polyQ-expanded ataxin-3 did not colocalize with the nucleolar marker protein B23 (Fig. 4B). To further confirm the nuclear localization of the arginine-containing RAN proteins, we employed another nucleolar marker protein, nucleolin (NCL). The colocalization of poly-GRs and the nucleolar markers was clearly observed using confocal microscopy (Fig. 4C and D). In addition, poly-PR localized to the nucleoli (data not shown), indicating that the nuclear inclusion bodies formed by poly-GR and poly-PR include nucleolar arginine-enriched RAN proteins but not protein aggregates.

poly-GR and poly-PR, but not other RAN proteins, specifically induce B23 translocation and nucleolar stress

To elucidate the biological consequences of the selective nucleolar localization of poly-GR and poly-PR, we examined the cells at the time when poly-GR and poly-PR began to kill the cells. Both RAN proteins tended to display an exclusively nuclear distribution from 24 to 48 h post-transfection (Fig. 5A). Moreover, during that period, the poly-GR and poly-PR-transfected cells exhibited nucleolar swelling and increased nucleoplasmic translocation of B23, a nucleolar stress sensor and a marker of nucleolar stress that exhibits protective properties (33,34) (Fig. 5B and C).

To gain further insight into the abnormal nucleolar changes mediated by arginine-enriched RAN proteins, we employed mass spectrometry-based proteomic analysis to identify the binding partners of the arginine-enriched RAN proteins, revealing an interactome that included 40S/60S ribosomal proteins, heterogeneous nuclear ribonucleoproteins (hnRNPs), B23 and 14 other proteins that strongly interacted with synthetic poly-GR (Fig. 5E). Among these proteins, most were nucleolar proteins or RNA-binding proteins associated with RNA granules. Moreover,
Arginine-enriched RAN proteins form nuclear ‘aggregate’-like inclusions. (A) HEK293 cells were transfected with GFP, GFP-GAggggcc, GFP-GAggagca, GFP-GRggggcc, GFP-GRggaaga, GFP-GPggggcc, GFP-PAccccgg or GFP-PRccccgg. Twenty-four hours later, the cells were stained with Hoechst (blue) to visualize the nuclei. Subsequently, live cell imaging was performed using fluorescence microscopy. (B) The percentages of the cells containing ‘aggregates’ in A are quantified. (C) The transfected cells in A with RAN protein diffusely distribute in the cytoplasm and nucleus were quantified. (D) NSC-34 cells were transfected with GFP-tagged GRggggcc or GRggaaga. Then, the cells were fixed 24 h post-transfection and stained with DAPI. (E) SH-SY5Y cells were transfected with GFP, GFP-GAggggcc, GFP-GAggagca, GFP-GRggggcc, GFP-GRggaaga or GFP-PRccccgg. Twenty-four hours later, the cells were subjected to microscopic observation. The scale bars represent 10 μm.
Figure 3. None of the five RAN proteins or the nuclear ‘aggregates’ formed by poly-GR and poly-PR was associated with UPS or autophagy. (A) HEK293 cells were transfected with various GFP-tagged hexanucleotide repeat constructs. Twenty-four hours later, the cells were incubated in 100 nM Baf A1 and 10 μM MG132 for 12 h. Then, the cells were fixed and stained with DAPI. The scale bars represent 10 μm. (B) HEK293 cells were transfected and treated with the inhibitors described in A for 12 h. Then, cell lysates were collected and subjected to immunoblot using the indicated antibodies. The protein levels from three independent experiments were quantified (lower panel). The quantitative data are indicated as the means ± SEM, NS, no significance. (C) HEK293 cells were transfected with GFP-tagged Htt-60Q. Twenty-four hours later, the cells were treated with DMSO or MG132 for 12 h. Then, the cells were observed using fluorescence microscopy, and the numbers of Htt-60Q aggregates were quantified. The quantitative data are indicated as the means ± SEM, **P < 0.01. The scale bars represent 10 μm. (D) GFP-GRggggcc, GFP-GRggaga or GFP-PRccccgg was transfected into ATG5 wild type (ATG5 WT) or ATG5 knockout (ATG5−−/−) MEFs. Twenty-four hours later, the cells were visualized using fluorescence microscopy. The scale bars represent 10 μm.
B23 co-immunoprecipitated with poly-GR and poly-PR (Fig. 5D, and Supplementary Material, Fig. S5), further confirming that the latter two factors are two nucleolar proteins that are associated with essential components in the nucleolus, such as B23. Taken together, our data suggest that poly-GR and poly-PR may cause morphological changes in the nucleolus and the translocation of B23 from the nucleolus to the nucleoplasm potentially by directly associating with nucleolar components.

The nucleolus is the site where ribosomal RNA is synthesized and the nascent ribosome is assembled, and B23 directly regulates ribosome biogenesis and pre-ribosomal particle transport (33,35). We hypothesized that poly-GR- and poly-PR-mediated B23 translocation might lead to defects in nucleolar function and ribosomal RNA biogenesis. To address this issue, we evaluated the maturation 18S and 28S rRNA, which are processed from the precursor 45S rRNA in RAN protein-expressing cells. The qRT-PCR data showed that both the 18S and 28S rRNA levels were remarkably decreased following the expression of poly-GR or poly-PR (Fig. 6A). In contrast, poly-GA expression increased the 18S and 28S rRNA levels (Fig. 6A). Similar results were obtained using NSC-34 cells (data not shown).

Figure 4. The nuclear inclusions formed by GR and PR are localized to the nucleolus. (A) HEK293 cells were transfected with FLAG-tagged GRggggccc or GRggaaga. Twenty-four hours later, the cells were observed under PH microscopy. Arrowheads indicate the nucleoli. (B) HEK293 cells were transfected with GFP-tagged full-length or truncated ataxin-3. Twenty-four hours later, the cells were subjected to immunofluorescence staining using an anti-B23 antibody (red) and stained with DAPI. Arrowheads indicate the aggregates formed by polyQ proteins. (C) HEK cells were transfected as described in A. Twenty-four hours later, the cells were fixed and stained with anti-FLAG (red) and anti-NCL antibodies (green). (D) HEK293 cells were transfected with GFP-GRggggccc or GFP-GRggaaga. Twenty-four hours later, the cells were subjected to immunofluorescence using an anti-B23 antibody (red) and stained with DAPI. The scale bars indicate 10 μm.
Figure 5. Poly-GR and poly-PR, but not other RAN proteins, specifically induce B23 translocation from the nucleolus to the nucleoplasm. (A) GFP-GRggggcc, GFP-GRggaaga or GFP-PRccccgg was transfected into HEK293 cells. After 24 or 48 h, the cells were stained with Hoechst to visualize the nuclei. The merged images show the colocalization of the GFP and Hoechst fluorescent signals. (B) HEK293 cells were transfected with GFP-GRggggcc, GFP-GRggaaga or GFP-PRccccgg. Forty-eight hours later, the cells were fixed and subjected to immunofluorescence using an anti-B23 antibody (red) and stained with DAPI. The scale bars indicate 10 μm. (C) HEK293 cells transfected with the indicated plasmid were fixed 48 h after transfection and stained with an anti-B23 antibody and DAPI. The size of the nucleoli, as indicated by B23 staining, from 50 transfected cells was analyzed. The quantitative data are expressed as the means ±SEM, *P < 0.05. (D) HEK293 cells were transfected as described in B. Forty-eight hours later, cell lysates were collected and subjected to immunoprecipitation using an anti-GFP antibody. Co-immunoprecipitated B23 was detected using an anti-B23 antibody. (E) HEK293 cells were transfected with GFP-GRggaaga. The GFP-immunoprecipitated samples were subjected to mass spectrometric analysis. The list contains 14 proteins that associated with GFP-GRggaaga.
To further elucidate the mechanism by which RAN proteins induce cytotoxicity, we treated cells with actinomycin D (ActD), a well-established chemical that induces nucleolar stress (36), to mimic nucleolar stress and compared the characteristics of ActD-treated cells with those of RAN product-transfected cells. ActD treatment strongly induced nucleolar segregation (a morphological change in the nucleolus that reflects a nucleolar functional defect), as shown in Supplementary Material, Figure S6 and Figure 6B that small nucleolar fragments (10–30 small dots double-labeled by B23 and NCL) were scattered across the nucleus after ActD treatment, and B23 translocation to the nucleoplasm (Supplementary Material, Fig. S6, and Fig. 6B) as previously reported (33), whereas the protein synthesis inhibitor cycloheximide (CHX) failed to induce nucleolar segregation or B23 translocation (Fig. 6B). Importantly, our results showed that poly-GR and poly-PR, but not poly-GA, poly-GP or poly-PA, strikingly...
induced nucleolar segregation and B23 translocation after an extended transfection period, similar to the results of ActD treatment (Fig. 6C and D). Furthermore, poly-GR and poly-PR expression in combination with ActD exerted the same effect on nucleolar segregation, B23 translocation and cell death as ActD treatment alone (Fig. 6C–F), indicating that the arginine-containing RAN proteins and ActD function in the same pathway that is closely associated with nucleolar stress. Taken together, our results identified nucleolar stress as a potential fundamental contributor to RAN translation-mediated C9ALS/FTD and revealed that this mechanism is specifically triggered by arginine-containing RAN proteins but not other RAN proteins.

Arginine-containing and non-arginine-containing RAN proteins exerted opposite effects on stress granule formation

Next, we examined whether RAN proteins affect stress granules, which are mRNA- and RNA-binding protein-enriched cytoplasmic compartments that associate with multiple proteins involved in ALS/FTD (37, 38). We transfected cells with the RAN constructs and analyzed stress granule formation by visualizing the classic stress granule marker protein G3BP1. Surprisingly and interestingly, our results showed that poly-GA, poly-GP and poly-PA induced stress granule formation under normal conditions (Fig. 7A and D), whereas poly-GR and poly-PR, as well as ActD, strikingly inhibited stress granule formation in cells treated with arsenite, a stress granule inducer (Fig. 7B, C, E and F), suggesting that a functional association between RAN proteins and stress granules may contribute to neurodegeneration.

Discussion

Although great efforts and rapid progress have been made to elucidate the mechanism of C9ALS/FTD in the 3 years because a mutation in the C9orf72 gene was identified as the predominant cause of ALS/FTD (39), whether the repeat expansion in the non-coding region of C9orf72 initiates neurodegeneration via a loss-of-function or a gain-of-function mechanism remains unclear. Our present findings, along with findings recently reported by McKnight, Issacs and Partridge groups, reveal a novel mechanism by which this hexanucleotide repeat expansion may cause disease: a gain-of-function mechanism caused by unconventionally translated arginine-containing RAN proteins, which are associated with nucleolar stress and impaired RNA and protein (RNP) homeostasis.

In the case of a loss-of-function mechanism, homozygous patients would be expected to develop a more severe clinical phenotype and a distinct disease spectrum. With this expectation in mind, the finding that a homozygous patient fell within the typical range of disease symptoms (40) suggests that a gain-of-function mechanism likely plays a primary role in C9ALS/FTD. In support of this concept, knockdown of C9orf72 transcripts, including mRNAs, using antisense oligonucleotides (ASOs) did not exert a toxic effect on iPS-derivated motor neurons (12, 41) but rather improved the gene expression profile of these cells (41). Moreover, reducing C9orf72 expression in the mouse nervous system did not induce neuropathological defects (14), suggesting that a gain-of-function, rather than a loss-of-function, mechanism might be the predominant cause of C9ALS/FTD.

In the case of a gain-of-function mechanism, a fundamental question underlying C9orf72 hexanucleotide repeat expansion-mediated neurodegeneration is whether it is caused by RNA- or protein-induced toxicity. Although several studies reported that the RNA-mediated toxicity of GGGGCC repeats may be associated with neurodegeneration, given that ‘pure’ GGGGCC RNA repeats lacking an ATG start codon may continue to produce RAN proteins in cells (16, 23), it remains possible that the RAN proteins actually contributed to the neurotoxicity observed in those studies. Here, we successfully distinguished the potential toxicity caused by hexanucleotide RNA repeats from that caused by RAN proteins using synthetic RAN proteins that lacked the potential for RNA toxicity, such as poly-GA, which was encoded by GGAAGA, and poly-GR, which was encoded by GGAAGA. In cells that were transfected in parallel, we found that the GGGGCC repeat alone (No ATG), poly-GA, poly-GR and poly-PA did not display strong toxicity under normal conditions, whereas both pathogenic poly-GA and synthetic poly-GR were strikingly toxic to the cells (Fig. 1). In general, poly-PR displayed greater toxicity than that of poly-GR, possibly due to the stronger nuclear localization of poly-PR (Figs 1 and 2). Although 30-repeat GGGGCC alone displayed very weak cytotoxicity after an extended period (Fig. 1), based on the results reported by Xu et al. that a 30-repeat GGGGCC sequence was sufficient to induce RAN translation (16), we propose that the weak damage induced by GGGGCC repeats alone in our study may have resulted from RAN proteins, such as poly-GA, which were produced by GGGGCC repeats. Taken together, our findings indicate that the poly-GR and poly-PR RAN proteins, rather than the hexanucleotide RNA repeats, might represent the primary toxic species responsible for the neurodegeneration observed during C9ALS/FTD. Consistent with our findings, Mizielinska et al. reported that poly-GA and poly-PA, as well as GGGGCC ‘RNA-only’ repeats containing stop codons, did not exhibit any toxicity in Drosophila (28). This recent study, together with ours, provides strong evidence that poly-GR and poly-PR products may play a greater role in C9ALS/FTD than do other RAN translation products, including poly-GA, poly-GR and poly-PA.

The next issue was whether RAN protein aggregates contribute to C9ALS/FTD. RAN translation of GGGGCC GGGCCC repeats leads to the formation of aggregates of RAN polypeptides in patients. Although the contribution of these ‘RAN aggregates’ to disease pathogenesis remains unclear, they may be closely associated with disease pathogenesis, similar to the role of polyQ proteins in polyQ disorders. Nevertheless, in our cell culture model, we did not detect the formation of classic protein aggregates of RAN proteins that resembled poly-Q-containing protein aggregates (Supplementary Material, Fig. S4A–D), even following treatment with a proteasome inhibitor or an autophagy inhibitor that are widely used to induce protein aggregation (Figs 3 and 4). Considering that the RAN proteins in patients may contain thousands of repeats, constructs that induce a higher expression level and longer repeat expansion, rather than the 30–60 repeats used in the current study, need to be employed to evaluate the aggregation and cytotoxicity of RAN proteins in the future. In addition, to further examine the role of RAN protein aggregation in human disease, in vivo experiments using rodent models need to be performed to determine whether RAN protein aggregation occurs in the long term and to verify whether it is a pre- or post-symptomatic event. However, notably, no correlation between poly-GA aggregation and the degree of neurodegeneration in C9ALS/FTD patients was detected (42). Furthermore, accumulated poly-GP remained in C9orf72 ALS iPS-derivated neurons after ASO treatment despite a reduction in neurotoxicity, suggesting that the aggregation of poly-GA and poly-GP may not be a predominant mediator of neurotoxicity in C9ALS/FTD (12). In addition, poly-G, poly-A and poly-P polypeptides were not
Figure 7. Arginine-containing and non-arginine-containing RAN proteins exerted opposite effects on stress granule formation. (A) HEK293 cells were transfected with GFP, GFP-CAgagggcc, GFP-CAggagca, GFP-GRgagggcc, GFP-GRggaaga, GFP-GPGagggcc, GFP-PCccccgg or GFP-PRccccgg. Forty-eight hours later, the cells were subjected to immunofluorescence using an antibody against G3BP1 (red) and stained with DAPI. (B) HEK293 cells were transfected as described in A. Forty-eight hours later, the cells were incubated in arsenite (0.5 mM) for 20 min, followed by immunofluorescence using an antibody against G3BP1 (red) and staining with DAPI. (C) GFP was transfected into HEK293 cells. Forty-eight hours later, the cells were treated with DMSO or ActD (2 μg/ml) for 4 h, followed by treatment with arsenite (0.5 mM) for 20 min. Subsequently, the cells were fixed, subjected to immunofluorescence using an antibody against G3BP1 (red), and stained with DAPI. (D–F) The percentages of transfected cells containing stress granules treated as in A–C were quantified. The results are expressed as the means ± SEM, **P < 0.01. The G3BP1 images display the combined G3BP1 and DAPI fluorescent signals. The scale bars represent 10 μm.
aggregate-prone proteins, as previously reported (43). Based on this evidence, the aggregation of most, if not all, pathogenic RAN proteins in patients may represent a downstream effect of the initial toxic effect-induced proteotoxic stress, and RAN protein aggregation may serve as a self-protection process to deposit toxic soluble oligomers, as has been reported for TDP-43 aggregation in ALS/FTD (44) and for polyQ aggregation in Huntington’s disease (45).

Interestingly and importantly, the interacnome from our mass spectrometry analysis of the synthetic poly-GR-interacting proteins included B23, hnRNPs [such as hnRNP A2/B1, which has been genetically linked to ALS (46)], 40S/60S ribosomal proteins and other proteins previously found in the nucleolus (47,48). Most importantly, these binding partners were previously demonstrated to associate with GGGGCC RNA repeats (12,15,17–20) (Fig. 5E). Moreover, the C9orf72-coding protein associates with hnRNP A1 and hnRNP A2/B1 (7), and Kwon et al. showed that poly-GR and poly-PR associate with hnRNP A2 (hnRNP A2/B1) (29), indicating a potential functional overlap between the RNP associates with hnRNP A1 and hnRNP A2/B1 (7), and Kwon et al. showed that poly-GR and poly-PR associate with hnRNP A2 (hnRNP A2/B1) (29), indicating a potential functional overlap between the

Our study revealed a pathological link between two RNP-enriched cellular compartments, the nucleolus and stress granules upon nucleolar stress induced by RAN proteins. Studies have shown that nucleolar stress is involved in many human disorders, such as cancer, Parkinson’s disease and Huntington’s disease (49–51). Furthermore, it has recently been reported that nucleolar stress, in association with NCL and B23 translocation, is a significant characteristic of C9orf72 ALS patient cells and GGGGCC repeat-expressing cells (18). Here, we further demonstrated that arginine-rich RAN proteins directly affect the master regulator of nucleolar function, B23, which in turn induced nucleolar stress and impaired stress granule formation (Figs 5–7). Given that stress granules are thought to act as dynamic sorting stations for mRNAs in response to stress stimuli, the overall function of stress granules is cellular quality control via the temporary storage of mRNAs, the repression of translation and the deceleration of cell growth during stress, which may help the cell to maintain survival via protein aggregation (52). If this process is blocked under stressful or pathogenic conditions, the cells might be more sensitive to toxicity. We aimed to elucidate the precise roles of five independent RAN proteins in stress granule formation. Based on our findings, stress granule formation was strikingly inhibited in poly-GR- and poly-PR-transfected cells and in ActD-treated cells. Alternatively, the other RAN proteins, which lacked arginine, exerted the opposite effect (Fig. 7). Similarly, poly-GR and poly-PR expression decreased ribosomal RNA biogenesis, whereas poly-GA expression increased it (Fig. 6A). According to the previous studies, rRNA maturation could be induced by cellular stress such as heat-stress and oxidative stress (53,54) and may serve as a source of nutrients (55). In our study, we found that poly-GA RAN protein could significantly enhance stress granule formation, which could also be enhanced under heat-stress and oxidative stress (56). We reasoned that poly-GA proteins could induce cellular stress, which is associated with stress granule formation. In this scenario, the 18S and 28S rRNA maturation may result from poly-GA-mediated cellular stress. Collectively, these previously unrecognized effects of arginine-containing and non-arginine-containing RAN products on ribosomal maturation and RNA granule formation may have important implications for our understanding of the fundamental mechanism underlying C9ALS/FTD and other neurodegenerative disorders.

In conclusion, our study suggests that the dysfunction of essential nucleolar proteins, such as B23, which is induced by RAN translation of the expanded hexanucleotide repeats in C9orf72, lead to upstream toxic nucleolar stress, the downstream dysfunction of RNP homeostasis and, ultimately, neurological damage (Fig. 8). These findings are helpful for both understanding the critical events underlying RAN translation-associated pathogenesis and developing effective treatments for C9ALS/FTD. Strategies to block RAN translation or to alleviate nucleolar stress, which may preserve nucleolar function in dysfunctional cells, could aid the development of novel and likely unexpected therapeutic targets.

**Figure 8.** A schematic model illustrating the role of poly-GR and poly-PR in C9orf72 RAN translation-induced cytotoxicity. The results of the present study suggest a simple model in which B23 translocation and nucleolar stress represent the primary molecular determinants of C9ALS/FTD.
Materials and Methods

Plasmid construction

HA-Ub, EGFP-p62, N htt-60Q-EGFP, EGFP-N1-ataxin-Q80, EGFP-N1-ataxin-Q130Q, EGFP-N1-ataxin-Q130QF, FLAG-B23 and EGFP-B23 plasmids were described previously (33,57-61).

The 30-repeat No-ATG-ggggcc-GFP sequence was kindly provided by Dr Peng Jin (Emory University, USA). First, using p3 × FLAG-myc-CMV-24 (Sigma) as a template, we constructed FLAG-1 and FLAG-2 plasmids by deleting the nucleotide sequence C or GC preceding the EcoRI site via site-directed mutagenesis using the MutanBEST kit (Takara) to produce three distinct reading frames after the insertion of the (ggaaga)30 repeats from FLAG-GRggaaga or FLAG-GAggagca, respectively.

Secondly, we constructed pcDNA3.1-2-(ggagca)30 or pcDNA3.1-2-(ggaaga)30 through cutting (ggagca)30 or (ggaaga)30 from pcDNA3.1-2-(ggagca)30 or pcDNA3.1-2-(ggaaga)30 using EcoRI/BamHI and then subcloned into the pEGFP-C1 vector to generate FLAG-GPggggcc, FLAG-GGggggcc and FLAG-RGggggcc, respectively.

To construct plasmids GFP-GPggggcc, GFP-GRggggcc and GFP-GAggggcc, the (ggggcc)30 repeats were excised from FLAG-GRggaaga or FLAG-GAggagca via digestion at EcoRI and BamHI to generate a TAG stop codon preceding the (ggaaga)30 repeats.

Finally, FLAG-GAggagca-long was generated by subcloning the (ggagca)30 repeats from pcDNA3.1-2-(ggagca)30 or pcDNA3.1-2-(ggaaga)30 into FLAG-GAggagca or FLAG-RGggaaga at EcoRV and BamHI sites. We termed these synthetic RAN constructs flag-myc-cmv-N3 through digested at EcoRI/BamHI sites.

Materials and Methods

Cell culture, transfection and chemicals

Human embryonic kidney 293 (HEK293), mouse motor neuron cell line NSC-34, ATG5 WT or KO MEF cells were cultured in Dulbecco’s modified Eagle’s medium (DMEM) containing 10% fetal bovine serum (FBS) with penicillin (100 U/ml) and streptomycin (100 μg/ml). SH-SY5Y cells were cultured in DMEM/F12 (1:1) medium containing 10% FBS with penicillin (100 U/ml) and streptomycin (100 μg/ml). The cells were transfected with Lipofectamine 2000 reagent (Invitrogen) at 30% confluence in DMEM without serum according to the manufacturer’s instructions. 4′,6-diamidino-2-phenylindole (DAPI), Hoechst 33342, thapsigargin (TG), CHX, actinomycin D (ActD), arsenite and propidium iodide (PI) were purchased from Sigma. MG132 was purchased from Calbiochem, and Bafilomycin A1 (Baf A1) was purchased from Fermentek.

PI staining assay and MTT assay

Cells were incubated at room temperature for 15 min and then the positive stained cells (dead cells) were observed using fluorescence microscopy (Olympus IX71). The dead cells were counted, and the quantitative data were shown. For MTT assay, the cells were washed with DMEM without phenol red and incubated with MTT (3-(4,5)-dimethylthiahiazo(-z-y1)-3,5-di-phenytetrazoliumromide, purchased from Sigma) at the concentration of 0.5 mg/ml in DMEM without phenol red. After 3 h, the media were discarded and the formazan crystals were dissolved in DMSO (dimethyl sulfoxide). The optical density was measured by a photometer at 570 nm, and background at 630 nm was subtracted. The quantitative data were normalized to control and the ratios are presented as means ± SE from three independent experiments.

Immunoblot analysis and antibodies

The cells were collected at time indicated after transfection and were lysed in cell lysis buffer (25 mM Tris–HCl (pH 7.6), 1% NP-40, 150 mM NaCl and 1% sodiumdeoxycholate) in the presence of a protease inhibitor cocktail (Roche). Then, the proteins were separated by SDS-PAGE and then transferred onto a polyvinylidene difluoride membrane (Millipore). Immunoblot analysis was performed with the following primary antibodies: monoclonal anti-Nucleophosmin/B23 antibody, polyclonal anti-nucleolin, anti-OPTN, anti-VCP, anti-TDP43 and G3BP1 antibodies were purchased from Proteintech; monoclonal anti-GAPDH antibody and monoclonal anti-FLAG antibody and monoclonal anti-GFP antibody were purchased from Sigma. Rabbit polyclonal antibody against GFP, anti-HA and anti-LC3 antibodies was described previously (59,60,63).

Immunofluorescence and live cell imaging

For immunofluorescence assay, NSC-34, SH-SY5Y or HEK293 cells were washed with pre-warmed PBS and fixed with 4% paraformaldehyde in PBS at room temperature for 5 min. Cells were
permeabilized with 0.2% Triton X-100 in PBS for 3 min and blocked for 20 min with 0.2% FBS in PBST. Cells were incubated with primary antibody diluted with PBST at room temperature for 3 h, washed lightly with PBST for 5 min. Then, secondary antibody was added and incubated for 1.5 h at room temperature. The following fluorophore-labeled secondary antibodies were used: Alexa Fluor 594 (red)-conjugated Affinipure Donkey Anti-mouse and Anti-rabbit IgG, as well as Alexa Fluor 488 (green)-conjugated Affinipure Donkey Anti-mouse and Anti-rabbit IgG (Invitrogen). Then, the cells were incubated with Alexa Fluor 488 (blue), and the live cells were incubated with Hoechst 33342 (blue) for 5 min to visualize the nuclei. Finally, the stained cells or live cells were observed using Zeiss LSM 710 confocal microscopy or fluorescence microscopy (Olympus IX71).

**Immunoprecipitation**

HEK293 cells transfected with the indicated plasmids were collected 48 h after transfection and lysed in cell lysis buffer. Cell pellet was discarded after centrifugation at 12 000g for 20 min at 4°C, and the supernatants were incubated with polyclonal anti-GFP antibody overnight at 4°C. After incubation, protein G Sepharose (Roche) was used for precipitation. The beads were washed three times with W1 buffer [50 mm Tris–HCl (pH 7.6), 1% NP-40, 150 mm NaCl and 1% sodium deoxycholate], two times with W2 buffer [25 mm Tris–HCl (pH 7.6), 0.1% NP-40, 500 mm NaCl and 1% sodium deoxycholate] and one time with W3 buffer [50 mm Tris–HCl (pH 7.6), 0.1% NP-40 and 1% sodium deoxycholate]. Then, the proteins were eluted with SDS sample buffer for immunoblot analysis.

**Mass spectrometry analysis**

HEK293 cells transfected with GFP or GFP-GRggaaga were lysed, and the cell lysates were immunoprecipitated with anti-GFP antibody and protein G Sepharose (Roche). The proteins were eluted with SDS sample buffer and separated on SDS-PAGE gel, and the gel was subjected to silver staining. Gel lanes were cut into slices, destained, reduced, alkylated and digested with trypsin. Then, the peptides were injected on to an LTQ-Orbitrap Elite (Thermo Fisher Scientific).

**Quantitative real-time PCR (qRT-PCR)**

Total RNA was extracted from cells using TRIzol Reagent (Invitrogen). Real-time PCR analysis was carried out for quantitative measurement of the target RNA abundance with SYBR Green Real-Time PCR Master Mix (Takara) in a PCR detection system (Applied Biosystems). The following primers were used: 5′- AATCC CATCACCATCTTCCA 3′ and 5′- TGGACTCACAGGTACTCA 3′ for human GAPDH; 5′- AGAGTAAACGGGTCGTC 3′ and 5′-GGGG TGCGGAGAAAGG 3′ for human 28 s rRNA; 5′- GATGGTAGTC GCGGTGCC 3′ and 5′- GCCTTGCCCTTCTTGG 3′ for human 18 s rRNA. The levels of 18 s and 28 s rRNA were normalized to GAPDH.

**Supplementary Material**

Supplementary Material is available at HMC online.

**Acknowledgements**

We thank Dr Arthur Horwich (Yale University) for critical comments and suggestions on our manuscript; Dr Mingyi Xie, Dr Xinguo Chen (Yale University) and Dr Shengdong Ke (The Rockefeller University) for helpful discussions and suggestions; Dr Peng Jin (Emory University) for providing us the plasmids with GGGGCC repeats; Dr Mian Wu (University of Science and Technology of China) for providing us the FLAG-B23 and EGFP-B23 plasmids; Dr Weidong Le (Baylor College of Medicine, USA) for providing us NSC-34 cells; Dr Chunshan Gui (Soochow University) for providing us HEK293 cells.

**Conflict of Interest statement.** None declared.

**Funding**

This work was supported in part by the National High-tech Research and Development program of China 973-projects (2011CB504102), the National Natural Sciences Foundation of China (Nos 31200803 and 31371072), Natural Science Foundation of Jiangsu Province (BK2012181), a Project Funded by Jiangsu Key Laboratory of Translational Research and Therapy for Neuro-Psycho-Diseases (BM2013003) and a Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

**References**


