

The 5'-Untranslated Region of the C9orf72 mRNA Exhibits a Phylogenetic Alignment to the Cis-Aconitase Iron-Responsive Element; Novel Therapies for Amyotrophic Lateral Sclerosis

Monica A. Lu¹, Susruthi Rajanala¹, Sohan V. Mikkilineni¹, Catherine M. Cahill¹, Robert Brown², James D. Berry³, Jack T. Rogers¹

¹Neurochemistry Laboratory, Department of Psychiatry-Neuroscience, Massachusetts General Hospital—Harvard Medical School, Charlestown, MA, USA

²Department of Neurology, University of Massachusetts Medical School, Worcester, MA, USA

³Department of Neurology, Massachusetts General Hospital—Harvard Medical School, Boston, MA, USA

Email: jack.rogers@mgh.harvard.edu

Received 9 November 2015; accepted 18 March 2016; published 21 March 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The hexanucleotide repeat mutation in the intron-1 of the chromosome 9 open reading frame (C9orf72) is a frequent cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). Altered RNA folding plays a role in ALS pathogenesis in two ways: non-ATG translation of the repeat can lead to aggregates of the known C9orf72 specific dipeptide polymer, whereas the repeat also can form neurotoxic RNA inclusions that dose-responsively kill motor neurons. We report the presence of a homology in the 5'untranslated region (UTR) of the messenger RNA encoding C9orf72 with the iron responsive elements (IRE) that control expression of iron-associated transcripts and predict that this RNA structure may iron-dependently regulate C9orf72 translation. We previously report altered serum ferritin levels track with severity of ALS in patients. Here, we conduct bioinformatics analyses to determine the secondary structure of the 5'UTR in C9orf72 mRNA and find it aligned with IREs in the human mitochondrial cis-aconitase and L and H-ferritin transcripts. Comparison of the role of RNA repeats in Friedreich's ataxia and fragile X mental retardation suggests the utility of RNA based therapies for treatment of ALS. Antisense oligonucleotides (ASO) have been reported to therapeutically target these GGGGCC repeats. At the same time, because the function of C9orf72 is unknown, knockdown strategies carry some risk of inducing or compounding haploinsufficiency. We propose, for consideration, an approach that may enhance its therapeutic dynamic range by increasing the 5'UTR driven translation of C9orf72 protein to compensate for any potential ALS-specific or ASO-induced haploinsufficiency.

How to cite this paper: Lu, M.A., *et al.* (2016) The 5'-Untranslated Region of the C9orf72 mRNA Exhibits a Phylogenetic Alignment to the Cis-Aconitase Iron-Responsive Element; Novel Therapies for Amyotrophic Lateral Sclerosis. *Neuroscience & Medicine*, 7, 15-26. <http://dx.doi.org/10.4236/nm.2016.71003>

Keywords

Amyotrophic Lateral Sclerosis (ALS), Iron-Responsive Element (IRE), C9orf72 mRNA, Mitochondrial Aconitase (mACO), Frontotemporal Dementia (FTD), Amyloid Precursor Protein (APP), HIV Trans-Activation Response Element (TAR), Antisense Oligonucleotides (ASO), Iron-Regulatory Proteins-1 and -2 (IRP1 and IRP2)

1. Introduction

Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) are both devastating neurological diseases. Amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disorder causing weakness, as well as other symptoms mediated by loss of upper and lower motor neurons. Lower motor neuron degeneration results in fasciculations and muscle atrophy, and upper motor neuron degeneration result in increased reflexes and spasticity. ALS is usually fatal due to respiratory paralysis within 1 to 5 years of symptom onset. The average age of onset is 55 to 65 years old and prevalence of ALS is 3 to 5/100,000 [1]. Frontotemporal dementia (FTD) is the second most common cause of pre-senile dementia after Alzheimer's disease, with the mean age of onset in the fifth to seventh decades of life [2]. Degeneration of the frontal and temporal lobes of the brain lead to progressive changes in behavior, personality and language.

ALS and FTD are part of a spectrum of neurodegenerative disorders that share pathological and molecular features. 4% - 14% of ALS patients have FTD and up to 40% of ALS patients show mild cognitive changes with further testing [3]; in addition, FTD patients may develop motor neuron dysfunction symptoms [4]. Furthermore, trans-active response DNA binding protein 43 kDa (TDP-43) has been identified as a common pathological protein, seen in the majority of ALS cases and in FTD [5]. Large expansions of a GGGGCC hexanucleotide repeat in the first intron of the chromosome 9 gene open reading frame (C9orf72) are a frequent cause of both ALS and FTD [6].

Trinucleotide repeats are observed as a driving cause of other neurodegenerative diseases such as Huntington's disease [7] and Friedrich's ataxia [8]. Repeat expansions can cause disease through multiple mechanisms. Fragile X syndrome, for one, stems from haploinsufficiency due to a CGG repeat in the 5' untranslated region that inhibits translation of the essential Fragile X Mental Retardation protein [9]. While the exact mechanism mediating C9orf72-ALS is unknown, a great deal has come to be known about the gene, the repeat expansion, and its effects. Herein, we use a bioinformatics approach to delineate a previously unreported potential role of C9orf72 in iron homeostasis, which can be disrupted by the hexanucleotide expansion.

1.1. RNA Binding Proteins in ALS

Mutations in the genes encoding TAR DNA Binding Protein (TDP-43) and Fused in Sarcoma/Translated in Sarcoma (FUS/TLS) cause ALS [10] [11]. Both these RNA binding proteins are involved in gene expression regulation by transcription, RNA splicing, RNA transport, and translation. The involvement of TDP-43 and FUS/TLS in ALS and other neurodegenerative diseases emphasizes the role of altered RNA processing in neurodegeneration. This has led to the hypothesis that disrupted RNA interactions might underlie C9orf72-mediated ALS, a hypothesis supported by a substantial body of evidence.

1.2. C9orf72 Pathology

The hexanucleotide repeat expansion in the first intron of the C9orf72 gene has been identified as the most common mutation in familial ALS. C9orf72 has two main variants, which, through alternative splicing, lead to different isoforms of the same C9orf72 protein [6]. In transcript variant 1, the GGGGCC repeat is located in the upstream regulatory region. Transcript variant 2 is the focus of this bioinformatics paper; in transcript variant 2, the GGGGCC repeat is located in the first intron between two noncoding alternatively spliced first exons [12]. Previous bioinformatic studies showed C9orf72 has structural homology to Differentially Expressed in Normal and Neoplasia (DENN) proteins that are critical in G protein signaling and may be essential for survival. There-

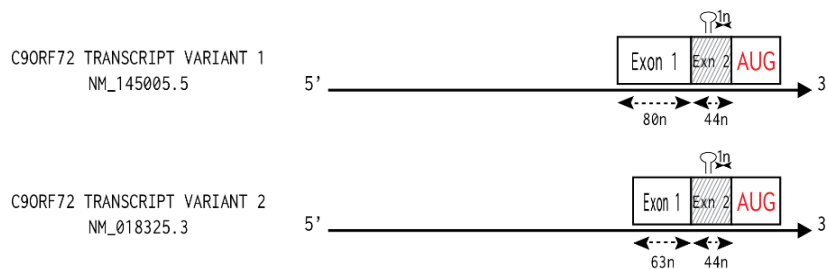
fore, DENN-like 72 (DENNL72) has been suggested as an alternative name for the C9orf72 gene [13].

Indeed, a subset of C9orf72 gene knockout studies confirmed that loss of function of C9orf72 causes motor deficits in a zebrafish model of ALS [14]. On the other hand, sustained antisense oligonucleotide (ASO)-mediated lowering of C9orf72 mRNAs provided a therapeutic impact from the RNA inclusions, which are a major hallmark of ALS pathology [15]. ASO treatment was well tolerated in mice producing no behavioral or pathological features characteristic of ALS/FTD and only caused limited RNA expression alterations. Thus, this therapy may well move into humans in relatively short order, making it critical that we learn the biological function of C9orf72 to better understand all of the potential effects of these gene-targeted therapies.

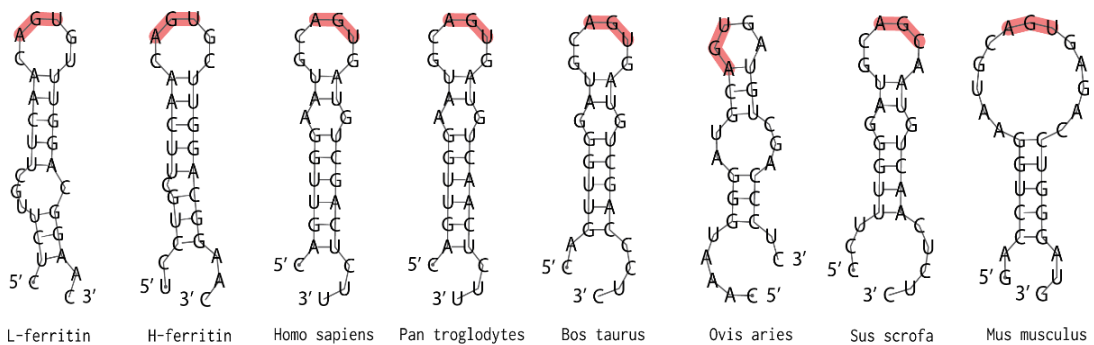
1.3. C9orf72 and Translational Control Circuits in Iron Homeostasis

In this bioinformatic review we will emphasize the organizational and sequence homology existing between the C9orf72 5'untranslated region (UTR) and the iron-responsive element (IRE) RNA stem loop in the 5'UTR of mitochondrial cis-aconitase and also in the mRNAs encoding L- and H-ferritin (Figure 1). Iron homeostasis is

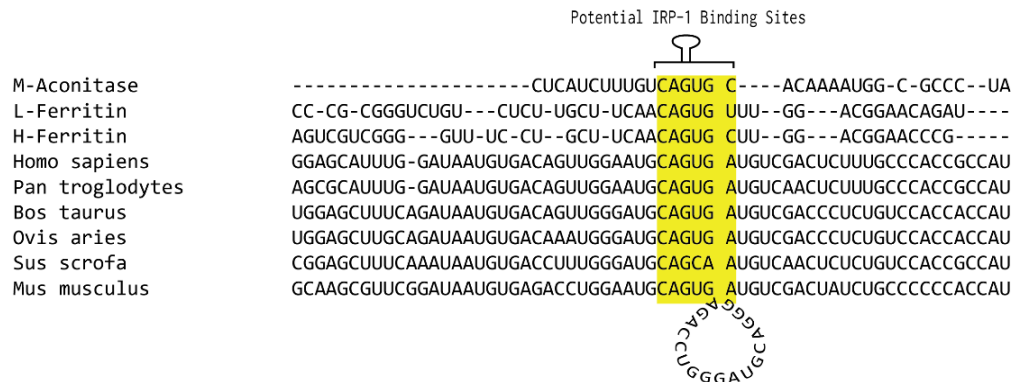
A. C9ORF72 Human 5' UTR Layout



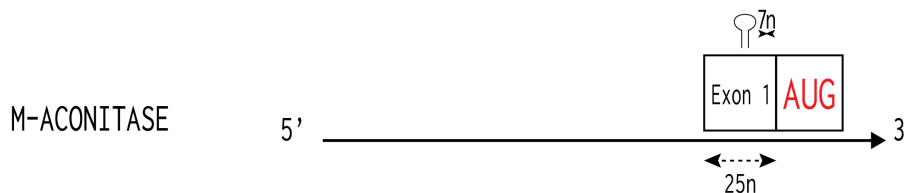
B. C9ORF72 5' UTR RNA Structures around IRE



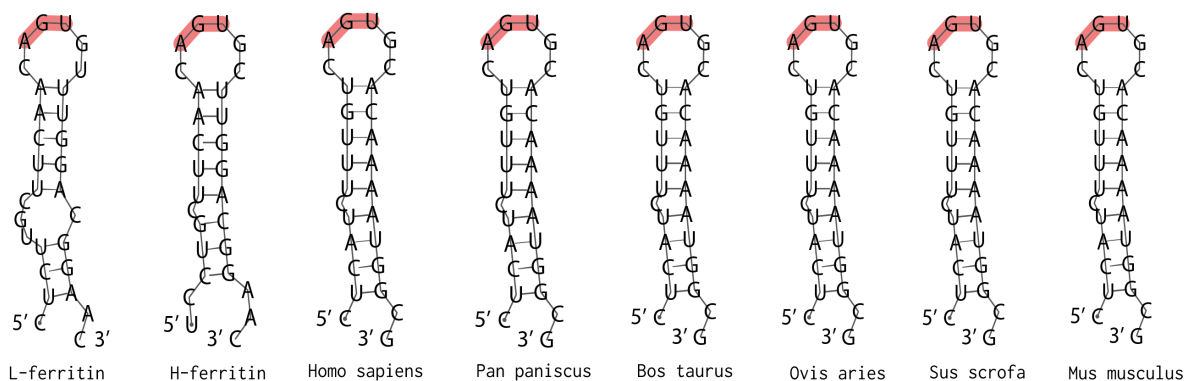
C. C9ORF72 Sequence Comparison to M-Aconitase and Across Species



D. M-Aconitase Homo sapiens 5'UTR Layout



E. M-Aconitase 5'UTR Structures around IRE



F. M-Aconitase Sequence Comparison Across Species

Potential IRP-1 Binding Sites

L-Ferritin	CUGUCUCUUGCUUC-AA CAGUGU UUG-GACGGAACAGAU---CCGG-GG--ACU-CUCUU
H-Ferritin	--GUUCCUGCUUC-AA CAGUGC UUG-GACGGAACCCG---GCGCUCGUUCC-CACC-
Homo sapiens	-----CU-CAUCUUUGU CAGUGC ACAAAAUGGCGCCUACAGCCUACUGGUGACUCGGCU
Pan paniscus	UGCGACCUCAUCUUUGU CAGUGC ACAAAAUGGCGCCUACAGCCUACUGGUGACUCGGCU
Bos taurus	AGCGACCUCAUCUUUGU CAGUGC ACAAAAUGGCGCCUACAGCCUACUGGUGUCCCGGCU
Ovis aries	AGCGACCUCAUCUUUGU CAGUGC ACAAAAUGGCGCCUACAGCCUACUGGUGUCCCGGCU
Sus scrofa	AGCGACCUCAUCUUUGU CAGUGC ACAAAAUGGCGCCUACAGCCUACUGGUGACCCGGCU
Mus musculus	UGCGACCUCAUCUUUGU CAGUGC ACAAAAUGGCGCCUACAGCCUACUGGUGACCCGGCU

Figure 1. Bioinformatics predictions for the C9orf72 IRE and mitochondrial aconitase. (A) The Layout of the 5'UTR specific RNA structures and the start codon in relation to the putative IRE in the C9orf72 mRNA, with two transcript variants numbered according to the NCBI RefSeq transcript collection. (B) RNA/FOLD computer program predictions of RNA stem loops from L-ferritin, H-ferritin, and C9orf72 mRNAs, including C9orf72 across mammalian species: Pan troglodytes, Bos taurus, Ovis aries, Sus scrofa, and Mus musculus. The human C9orf72 stem loop resembles the classical IRE RNA stem loop (5'CAGUGN3' loop motif) that controls iron-dependent L- and H-ferritin translation and transferrin receptor mRNA stability. The pseudotriloop AGU is highlighted at the apex of the L-ferritin and H-ferritin IREs, while the analogous AGU from C9orf72 species mRNAs is shown. (C) The 5'untranslated region (5'UTR) of the ALS-linked C9orf72 exhibits homology with L-ferritin, H-ferritin, and mitochondrial aconitase centered at its 5'CAGUGN3' loop motif, as well as evolutionary conservation across mammalian species (Percent phylogenetic homology to Homo sapiens over the core IRE region: Pan troglodytes 98.0%, Bos taurus 92.2%, Ovis aries 86.3%, Sus scrofa 82.4%, Mus musculus 86.3%). (D) The arrangement of 5'UTR structures and the start codon in relation to the IRE in the mitochondrial aconitase mRNA. (E) RNA/FOLD computer program predictions of RNA stem loops from L-ferritin, H-ferritin, and mitochondrial aconitase mRNAs, including m-aconitase across mammalian species: Pan paniscus, Bos taurus, Ovis aries, Sus scrofa, and Mus musculus. (F) The 5'untranslated region (5'UTR) of m-aconitase exhibits homology with L-ferritin and H-ferritin, as well as evolutionary conservation across mammalian species.

regulated by the iron-regulatory proteins IRP1 and IRP2. These RNA binding proteins post-transcriptionally regulate messenger RNA (mRNA) stem loops called iron-responsive elements (IREs). IRP1 and 2 bind to IREs with high affinity when there is an iron deficit and repress IRE-dependent translation [16] [17]. The two are less likely to bind IREs in an abundance of iron [18]. Iron influx releases iron regulatory proteins from the Light (L)- and Heavy (H)-ferritin transcripts to induce L- and H-chain translation [19]. Based on homologies in the 5'UTR, we hypothesize that a similar mechanism could influence the expression of C9orf72. Interestingly, ferritin levels are elevated in people with ALS.

1.4. A Putative Iron Responsive Element (IRE) in C9orf72

In the spectrum of IRE dependent regulation, human mitochondrial aconitase (mACO) mRNA is 30 fold less avidly bound to and regulated by iron regulatory proteins than the ferritin transcript [20]. In this report, we provide evidence that the 5'UTR of mRNA encoding C9orf72 exhibits a similar structure and location conservation to the IRE RNA stem loop in the mRNA encoding mACO (**Figure 1**). In the model based on this predicted 5'UTR structure, we show that the IRE-like RNA structure of C9orf72 transcript is strategically positioned in a similar manner to that found in human mACO mRNA in front of its start codon (**Figure 1**). These RNA stem loops can be readily tested for their capacity to post-transcriptionally regulate expression of this key ALS-linked gene in response to iron homeostasis.

Our bioinformatics alignments in the accompanying **Figure 1** will show that the position of the C9orf72 AUG start codon recapitulates the IRE arrangement for the mACO mRNA, one of the prototypical examples of an IRE, and may provide insight into a role of C9orf72 in iron homeostasis. IRP1, itself, can function as either active aconitase or an active RNA binding protein [21]. Mitochondrial dysfunction and oxidative damage play a role in neurodegenerative diseases and aging. Aconitase is affected by Huntington's disease and Friedreich's ataxia [22]. ALS may also affect aconitase levels. Protein aggregation is a proposed mechanism for the pathogenesis of ALS, and in one study, in the spinal cord tissues of sporadic ALS patients, aconitase levels were enhanced in these patients compared to the controls [23].

2. Materials and Methods

RNA sequences from the C9orf72 gene were located using the NCBI Gene database (NM_145005.5 variant-1 and NM_018325 variant-2). Since the 5'UTRs were of primary interest the coding regions were disregarded, apart from the initiating AUG [24]. For the C9orf72 transcript, the splice junction between the first two exons occurred 44 nt upstream from the conserved CAGUG, which in turn is immediately adjacent to the AUG start codon. As shown, there is high degree of sequence conservation in the second exon of the C9orf72 mRNAs among the other species investigated (**Figure 1(C)**). In order to study a balanced sequence, 25 nucleotides on each side of the putative IRP1 binding site (5'CAGUG3') were used as a start-point to create 50 nucleotide RNA query sequences (mouse and rat had 52 nucleotide sequences due to insertions). RNA sequences were aligned using the ClustalX2 graphical program to find evolutionary conservation between species. The AUG start region of the CDS was then used as a reference point to align the sequences, allowing a comparison of the sequences in both exons centered on the splice junction [18].

Secondary structure folding of these RNA sequences was generated by the RNA-Fold web server at the University of Vienna and was annotated using the RNA-Fold software package utilities. In this report, we used the RNA-Fold server in a manner that previously provided the most probable RNA secondary structure for SNCA 5'UTR sequences based on minimum free energy calculations [24] [25]. The alignment homology were calculated, as we described [24], by comparing the C9orf72 RNA sequence from different species against the H. sapiens sequence. To emphasize specificity, we routinely studied antisense versions of the C9orf72 5'UTR and found these RNA-fold predictions did not generate the RNA stem loop with the "AGU" tri-loop characteristic of the canonical iron-responsive element.

We calculated the alignment homology by comparing the different species RNA sequences (as listed in **Figure 1(C)**) against the Homo sapiens sequence of C9orf72 mRNA. We also calculated the homology between the core L- and H-ferritin IREs with that of C9orf72 and APP mRNAs. Only nucleotides that matched respective to the Homo sapiens sequence were scored a point; we determined the percent homology on each side of the AGU tri loop (putative core IRP1 binding site) by totaling the points scored and dividing by the total number of nucleotides on that side. The results from each side were then compared to illuminate the difference in conservation.

In sum, the percentage nucleotides that matched to the Homo sapiens sequence were scored from the total nucleotide positions on each side of a core CAGUGN IRE motif as observed for ferritin L- and H' and cis-aconitase transcripts. The results were then compared to better understand the difference in conservation immediately in front of the start codon.

3. Results

We conducted bioinformatics sequence comparisons and RNA secondary structure predictions to determine the secondary structure of the 5'UTR of the transcript encoding C9orf72 relative to the IREs encoded by mitochondrial aconitase, H-ferritin, and L-ferritin mRNAs. We determined the minimum energy of RNA folding for the 5'UTRs sequence of C9orf72 (Gibbs free energy = -4.64 kcal/mol), redrawn in **Figure 1(B)**, and the predicted structures for C9orf72 were compared with the range of known canonical IREs from H-ferritin and L-ferritin mRNAs to the IRE in the 5'UTR of mitochondrial aconitase mRNAs from different species.

Our alignment revealed the 5'UTR of C9orf72 mRNA encoded a core 5'-CAGUG-3' sequence that aligned with the IRE sequences encoded in the mRNAs for the ferritin H- and L-chains and mitochondrial aconitase. Whereas surrounding sequences diverged, homology was most prominent in proximity to the canonical 5'-CAGUG-3' (**Figure 1(C)**). Specifically, there is 80% homology between the putative IRE in C9orf72 mRNA and that in cis-aconitase mRNA, differing by only one base in the loop region. We ensured specificity in alignment and RNA stem loop studies to clarify the absence of an Iron responsive Element homology in the control 5'UTRs for **b-actin** and **b-tubulin** mRNAs equivalent sites registered the same distance from their start codons [26]. Of interest, there are several reports documenting a high affinity RNA binding interaction between the conserved AGU tri-loop in the ferritin and aconitase mRNA IRE RNA stem loops and the active site in Iron-regulatory-protein-1 (IRP1), as has been established by mutational studies and X-Ray crystallography [27]. In **Figure 1**, we observed that the apical AGU site in C9orf72 mRNA and in aconitase mRNA were, in each case, on the apex of these related RNA stem loops in conserved positions in front of their respective AUG codons. The homology is depicted from the colored lettering in the alignments (**Figure 1(B)**, **Figure 1(E)**). The location of the apical AGU tri-loop in the C9orf72 RNA stem loop was one nucleotide from the AUG start site located in exon 2. This is downstream of the ALS mutation repeat in the first intron of transcript variant 2 (**Figure 2**). The proximity of the IRE-like stem loop in C9orf72 to its start codon was found to be analogous to the AGU tri-loop in the functional IRE of the mRNA encoding mitochondrial aconitase, which is located 7 nucleotides from the AUG start site (**Figure 1(F)**).

Across mammalian species, C9orf72 specific 5'UTR sequences showed strong evolutionary conservation, and RNA secondary structure predictions were generated for Pan troglodytes, Bos taurus, Ovis aries, Sus scrofa, and Mus musculus. We identified a significant homology within the C9orf72 5'UTR across species (percent homology to Homo sapiens over the core IRE region: Pan troglodytes 98.0%, Bos taurus 92.2%, Ovis aries 86.3%, Sus scrofa 82.4%, Mus musculus 86.3%) (**Figure 1(C)**). The putative IRP1/IRP2 binding site is a conserved RNA stem loop that is depicted in an apical CAGUG loop in all the mammalian species.

4. Discussion

4.1. Biological Implications for the Presence of an IRE-Like RNA Stem-Loop in the C9orf72 Transcript

In this study, we identified a potential IRE-like RNA structure in C9orf72 mRNA and demonstrated that it is located in the exon following the hexanucleotide repeat expansion that causes ALS and FTD. Thus, C9orf72 may well be regulated by iron-dependent post-transcriptional control, via IRP1 binding.

We noted critical features about the 5'UTR specific RNA stem loop immediately in front of the start codon in the C9orf72 mRNA. First, there are three differentially phased AUG start codons in the 5'UTR of C9orf72 mRNA that are evolutionarily conserved across all of the mammalian species with the third being the functional start codon (**Figure 1(C)**). Second, the first AUG is positioned 22 nucleotides upstream from the start site and the second potential start codon sequence is located 8 nucleotides from the start site. Third, we noted the mini-open reading frame for the potential methionine (ATG)-glutamine (CAG) dipeptide in the putative IRE region after the second AUG in the main variant of the C9orf72 transcript, which warrants further investigation. This small open reading frame exists with an in-phase opal stop codon (UGA) downstream of the second start

C9orf72 - Potential Therapies

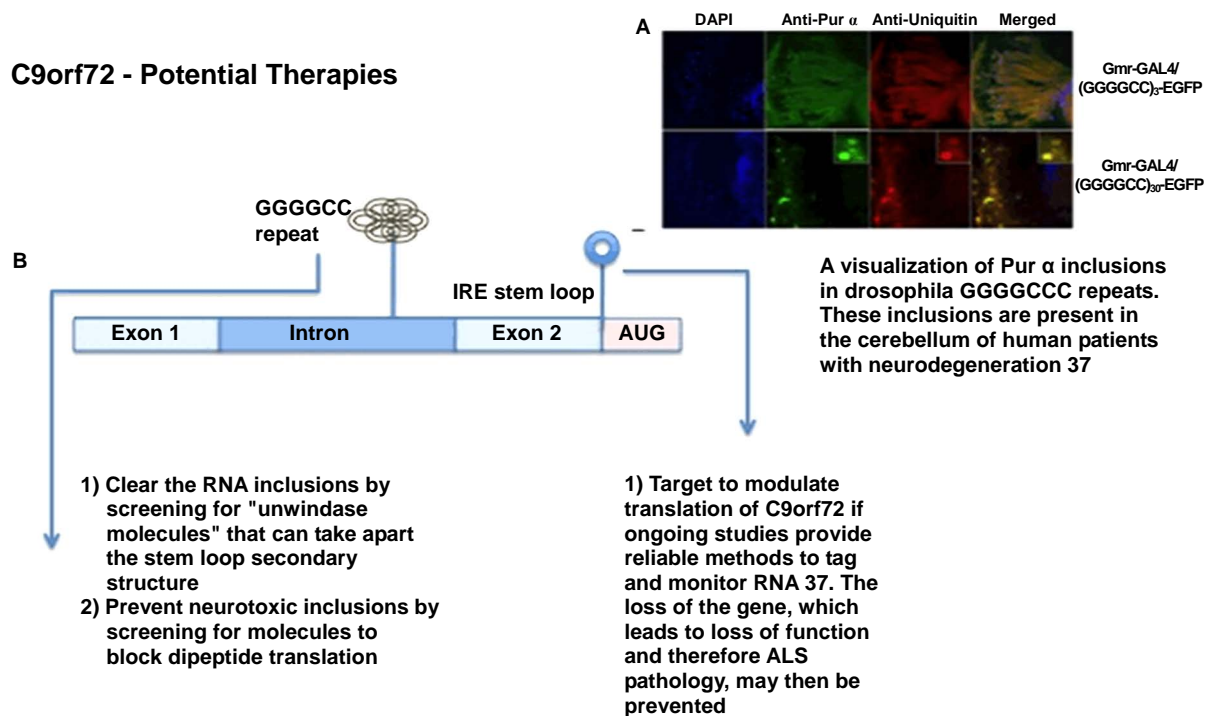


Figure 2. Potential ALS therapies targeting C9orf72. (A) A visualization of RNA inclusions in drosophila and human when expressing GGGGCC repeats in the C9orf72 gene. These inclusions have been stained with antibodies and are present in the cerebellum of patients with neurodegeneration as a result of ALS and FTD. (B) A cartoon of the C9orf72 gene showing the location of the RNA secondary stem loop on the intron between the two first exons. The IRE-like structure is also present upstream of the AUG start codon. Therapies targeting this stem loop as well as the IRE loop are detailed.

codon. The possibility that a natural dipeptide may be translated from wild type C9orf72 warrants further investigation since Mori *et al.*, 2013 reported dipeptides may be translated through repeat-associated non-ATG translation (RAN translation) as a cause of ALS as in other trinucleotide diseases. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTL/ALS.

4.2. Iron, Inflammation and ALS: The Precedent of Translational Control of the Amyloid Precursor Protein (APP)

Iron dys-homeostasis and inflammation are both associated with increased risk for ALS via many pathways that also occur for other neurodegenerative diseases such as Alzheimer's disease (AD) and Parkinson's disease [28]. Consistent with an iron link to accelerate ALS pathology, we reported elevated serum ferritin levels in ALS patients [28]. Increased ferritin levels may reflect an iron imbalance or inflammation, which has been demonstrated in ALS [29]. Events in motor neurons leading to local iron deficit (micro-anemic-events) could, for example, down-regulate Iron regulatory protein-1/-2-dependent C9orf72 translation via its IRE RNA stem loop shown in **Figure 1**. This scenario would cause local C9orf72 haploinsufficiency in affected motor neurons. On the other hand, excess iron would be predicted to increase C9orf72 levels with unknown pathological consequences to ALS progression.

As a precedent, it is informative to consider how inflammation may accelerate the course of AD in part by modulating the rate of translation of the amyloid precursor protein. An Interleukin-1 (IL-1)-responsive RNA domain, the "acute box" motif, is present immediately downstream from the IRE and upstream of the start codon in the amyloid precursor protein transcript (APP mRNA similar to L- and H-mRNAs of the iron storage protein ferritin) [26] [30]. This is relevant to IL-1 induction of APP translation and brain amyloid buildup as a component of the progression of AD during an accompanying inflammatory cascade in the brain cortex [31] [32]. Inflammatory IL-1 also increases APP translation by directly modulating the phosphorylation of IRP-1/-2 [33] to consequently release IRP-1/IRP2 from inhibiting translation (resulting in neurotoxic amyloid buildup) [19].

We can thus compare the APP model of translational control as a precedent to consider the impact of IL-1 and iron on C9orf72 gene expression at the translational level as a pathological pathway to potentially accelerate the course of ALS. When ascribing how inflammation may modulate rates of translational control of C9orf72 expression, we note that there is not sufficient coding capacity for the existence of IL-1 acute box RNA stem loop between the novel IRE that we identified and the AUG start codon in C9orf72 mRNA (**Figure 1**). Therefore, in the absence of an acute box in the 5'UTR in C9orf72 mRNA, inflammatory cytokines (IL-1) that are known to progress ALS would be predicted to only be capable of modulating rates of C9 translation by targeted phosphorylation or dephosphorylation of inducible IRP1-IRP2 as translation repressors. Such an inflammation induced repression of C9orf72 translation could indeed theoretically cause haploinsufficiency of this essential ALS gene product and accelerate ALS progression via this pathway.

4.3. Therapeutic Implications (RNA Based Pathology of GGGGCC Repeats)

A number of mechanisms may cause the pathology seen in frontal temporal dementia (FTD) and ALS patients. FTD and ALS patients who bear the pathological C9orf72 repeat expansion can exhibit repeat-associated non-ATG translation (RAN translation) of the C9orf72 GGGGCC-repeat expansion, which has been reported to translate dipeptide repeat proteins [34]. These patients show the presence of dipeptide repeat aggregates in brain regions such as the hippocampus and frontotemporal neocortex as well as TDP-43 pathology [35] (**Figure 2(A)**).

Alternatively, the repeat-associated non-ATG translation may occur, as in other trinucleotide repeat diseases. Here, longer GGGGCC repeat expansions may generate toxic RNA foci [36]. The repeat GGGGCC expansion in mammalian cells and *Drosophila* systems have been shown to cause cell death, through an interaction between Pur α , a major protein that binds the repeat GGGGCC [37]. These findings suggest that the repeat GGGGCC expansion seen in ALS and FTD patients may sequester RNA binding proteins from their normal function, resulting in neurodegeneration.

The GGGGCC repeat seen in C9orf72 can also form RNA G-quadruplex structures. These G-quadruplex structures are highly stable nucleic acid structures and have been implicated in alternative splicing, RNA transcription, translation, and telomere stability. G-quadruplex complexes may also cause pathology, as repeat tract lengths positively correlate with disease severity in other repeat expansion diseases, like myotonic dystrophy type 1 (DM1) with repeats of CTG [38] [39]. Ribonuclear foci are formed, and it is thought that ALS-FTD patients may have a similar pathogenesis where mutant transcripts containing the GGGGCC expansions may aggregate in the nucleus and sequester binding proteins [40].

Studies also show that the RNA inclusions and peptides generated by non-AUG mediated translation of the hexanucleotide repeat may lead to gain-of-function [41]. In other neurodegenerative diseases such as Huntington's disease, a polyglutamine tract of a CAG RNA triplicate is seen that contributes to the pathology and may be subject to aggregation. There are several interesting parallels between ALS and Friedreich's ataxia (FRDA), a genetic disorder that results in mutations that limit the production of the protein frataxin. Similar to ALS, the most common cause of FRDA is a GAA trinucleotide repeat expansion in the first intron of the gene encoding frataxin. This repeat results in DNA structures that impair frataxin transcription leading to a deficiency of the protein. Frataxin deficiency then leads to a breakdown in cellular iron homeostasis [42]. Potential gene therapies targeting the frataxin promoter have been researched; these therapies increase the expression of frataxin to counteract the effects of the GAA repeat [43].

4.4. Therapeutic Targeting GGGGCC Repeats While Also Activating Translation Driven by the C9orf72 5'-Untranslated Region

We propose novel small molecule therapies for ALS and FTD based on the strategy of screening and identifying small molecules to activate translation of the specific IRE-like RNA stem loop in 5'UTR of C9orf72 mRNA to thus enhance its translation in front of the start codon. Our goal is to provide an important therapeutic step to restore or preserve C9orf72 translation, given that it may be a pro-survival protein. Antisense therapies are being developed to target the GGGGCC RNA repeats that potentially generate dual neurotoxicity via 1) RNA inclusions formed from RNA guanine quadruplexes (G-quadruplexes) and 2) RAN Translation dipeptides. Translational activation of the IRE-like RNA stem loop in the C9orf72 transcript might to compensate for the loss of C9orf72 protein when using antisense strategies. We also propose using the parallel use of RNA targeting strategies to activate compensatory C9orf72 translation to regain potential loss of a key metabolic protein.

We consider the multiple pathogenic pathways of the disease via RNA-based expression of C9orf72 in neurons. The following are the levels of RNA-based regulation of C9orf72 to be considered.

A. The RNA secondary structure of the repeat itself, leading to gain of function, may be targeted. The hexanucleotide repeat causes RNA G-quadruplex structures. These inclusions may be prevented by screening for small molecules that are capable of unwinding the RNA stem loop secondary structure to therefore clear the RNA inclusions from the cells (**Figure 2(A)**, **Figure 2(B)**). This has been achieved by use of antisense oligonucleotides. Of note, the use of the spinach technique is an endpoint assay [44] may permit small molecular targeting of these RNA inclusions and the G-quadruplexes.

B. The hexanucleotide G-quadruplex repeat also presents a target for ribosomal entry for dipeptide protein translation, and thus for increased production of such dipeptide aggregates in regions of the brain. This is presented in the bioinformatics scheme of the C9orf72 open reading frame shown in **Figure 2**. These seminal findings support ongoing screening for small molecule agents that block dipeptide translation from G-quadruplexes to prevent these neurotoxic inclusions in ALS patients (**Figure 2(B)**).

C. C9orf72 shows structural homology to DENN proteins, and DENN modules function as guanine nucleotide exchange factors for small GTPases [45]. Though the function of C9orf72 has not been completely elucidated, C9orf72 may regulate endosomal trafficking and autophagy in neuronal cells [46]. We propose an additional potential regulatory and functional role for C9orf72 pertinent to citric acid and iron metabolism. Indeed, relevant to this new bioinformatic link for the role of C9orf72 in metabolism, there is debate as to the whether the loss of C9orf72 causes motor neuron deficiency in organisms higher than zebrafish [4]. Thus, aspects of disease pathology of ALS and FTD can potentially result from haploinsufficiency, where even the lack of half of normal C9orf72 levels could compromise neuronal function.

5. Conclusions

It is clear that an RNA stem loop, as depicted in **Figure 1** in front of the start codon, should constitute an attractive pharmacological target for small molecules capable of increasing translation of C9orf72 mRNA. **Figure 2** presents our model that C9orf72 translation activators should compensate for the loss of this key neurodegenerative disease associated protein, inclusive of when conducting therapies with antisense oligonucleotides (ASO). It may be desirable to combine the use of such activators when employing ASOs to target the GGGGCC repeats to reduce RNA/dipeptide inclusions specific to C9orf72 to simultaneously stimulate increased translation of the C9orf72 protein open reading frame. We propose to compensate for reduced levels of this potentially key iron-associated protein when ALS patients receive ASO treatments.

The 5'untranslated region RNA structure depicted in **Figure 1** presents an attractive therapeutic RNA-based target for identifying small molecules translation activators of C9orf72 via gene-transfection based screens as they are previously conducted for the mRNAs for amyloid precursor protein, alpha-synuclein and the prion precursor protein. Current research presents a technique to tag and facilitate fluorescence for C9orf72 RNAs, as it has been achieved with proteins for Fragile X syndrome, namely the "spinach technique". This method can monitor the use of mainstream therapeutic approaches where antisense oligonucleotides are being used to potentially reduce ALS specific RNA inclusions and/or prevent translation of the toxic di-peptide repeats from the C9orf72 mRNA. Nevertheless co-treatments with small molecule C9orf72 translation activators of the RNA stem loop shown in **Figure 1** may represent a pharmacological means to avoid the excessive loss of C9orf72 protein in conditions when its transcript levels are reduced, an event that would have worsened ALS pathology [44]. Upcoming research will reveal how the GGGGCC hexanucleotide repeat causes the RNA inclusions and binds ribosomes, at which time we propose it will become clearer whether therapies to increase C9orf72 translation would prove an adjunct reliable cure to ALS.

References

- [1] Al-Chalabi, A. and Hardiman, O. (2013) The Epidemiology of ALS: A Conspiracy of Genes, Environment and Time. *Nature Reviews Neurology*, **9**, 617-628. <http://dx.doi.org/10.1038/nrneurol.2013.203>
- [2] Johnson, J.K., Diehl, J., Mendez, M.F., Neuhaus, J., Shapira, J.S., *et al.* (2005) Frontotemporal Lobar Degeneration: Demographic Characteristics of 353 Patients. *Archives of Neurology*, **62**, 925-930. <http://dx.doi.org/10.1001/archneur.62.6.925>
- [3] Phukan, J., Elamin, M., Bede, P., Jordan, N., Gallagher, L., *et al.* (2012) The Syndrome of Cognitive Impairment in

- Amyotrophic Lateral Sclerosis: A Population-Based Study. *Journal of Neurology, Neurosurgery & Psychiatry*, **83**, 102-108. <http://dx.doi.org/10.1136/jnnp-2011-300188>
- [4] Lomen-Hoerth, C., Anderson, T. and Miller, B. (2002) The Overlap of Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. *Neurology*, **59**, 1077-1079. <http://dx.doi.org/10.1212/WNL.59.7.1077>
- [5] Neumann, M., Sampathu1, D.M., Kwong1, L.K., Truax, A.C., Micsenyi1, M.C., *et al.* (2006) Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. *Science*, **314**, 130-133. <http://dx.doi.org/10.1126/science.1134108>
- [6] Renton, A.E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., *et al.* (2011) A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron*, **72**, 257-268. <http://dx.doi.org/10.1016/j.neuron.2011.09.010>
- [7] Tsoi, H. and H.Y. (2015) Chan, Expression of Expanded CAG Transcripts Triggers Nucleolar Stress in Huntington's Disease. *Cerebellum*, **12**, 310-312. <http://dx.doi.org/10.1007/s12311-012-0447-6>
- [8] Libri, V., Yandim, C., Athanasopoulos, S., Loyse, N., Natisvili, T., *et al.* (2015) Epigenetic and Neurological Effects and Safety of High-Dose Nicotinamide in Patients with Friedreich's Ataxia: An Exploratory, Open-Label, Dose-Escalation Study. *The Lancet*, **384**, 504-513. [http://dx.doi.org/10.1016/S0140-6736\(14\)60382-2](http://dx.doi.org/10.1016/S0140-6736(14)60382-2)
- [9] Santoro, M.R., Bray, S.M. and Warren, S.T. (2015) Molecular Mechanisms of Fragile X Syndrome: A Twenty-Year Perspective. *Annual Review of Pathology*, **7**, 219-245. <http://dx.doi.org/10.1146/annurev-pathol-011811-132457>
- [10] Kwiatkowski Jr., T.J., Bosco, D.A., LeClerc, A.L., Tamrazian, E., Vanderburg, C.R., *et al.* (2009) Mutations in the FUS/TLS Gene on Chromosome 16 Cause Familial Amyotrophic Lateral Sclerosis. *Science*, **323**, 1205-1208. <http://dx.doi.org/10.1126/science.1166066>
- [11] Van Deerlin, V.M., Leverenz, J.B., Bekris, L.M., *et al.* (2008) TARDBP Mutations in Amyotrophic Lateral Sclerosis with TDP-43 Neuropathology: A Genetic and Histopathological Analysis. *The Lancet Neurology*, **7**, 409-416. [http://dx.doi.org/10.1016/S1474-4422\(08\)70071-1](http://dx.doi.org/10.1016/S1474-4422(08)70071-1)
- [12] Cruts, M., Gijselinck, I., Van Langenhove, T., *et al.* (2013) Current Insights into the C9orf72 Repeat Expansion Diseases of the FTL/ALS Spectrum. *Trends in Neurosciences*, **36**, 450-459. <http://dx.doi.org/10.1016/j.tins.2013.04.010>
- [13] Levine, T.P., Daniels, R.D., Gatta, A.T., *et al.* (2013) The Product of C9orf72, a Gene Strongly Implicated in Neurodegeneration, Is Structurally Related to DENN Rab-GEFs. *Bioinformatics*, **29**, 499-503. <http://dx.doi.org/10.1093/bioinformatics/bts725>
- [14] Ciura, S., Lattante, S., Le, B.I., *et al.* (2013) Loss of Function of C9orf72 Causes Motor Deficits in a Zebrafish Model of Amyotrophic Lateral Sclerosis. *Annals of Neurology*, **74**, 180-187. <http://dx.doi.org/10.1002/ana.23946>
- [15] Lagier-Tourenne, C., Baughn, M., Rigo, F., *et al.* (2013) Targeted Degradation of Sense and Antisense C9orf72 RNA Foci as Therapy for ALS and Frontotemporal Degeneration. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, E4530-E4539. <http://dx.doi.org/10.1073/pnas.1318835110>
- [16] Bandyopadhyay, S. and Rogers, J.T. (2014) Alzheimer's Disease Therapeutics Targeted to the Control of Amyloid Precursor Protein Translation: Maintenance of Brain Iron Homeostasis. *Biochemical Pharmacology*, **88**, 486-494. <http://dx.doi.org/10.1016/j.bcp.2014.01.032>
- [17] Bandyopadhyay, S., Cahill, C., Balleidier, A., *et al.* (2013) Novel 5' Untranslated Region Directed Blockers of Iron-Regulatory Protein-1 Dependent Amyloid Precursor Protein Translation: Implications for Down Syndrome and Alzheimer's Disease. *PLoS ONE*, **8**, e65978. <http://dx.doi.org/10.1371/journal.pone.0065978>
- [18] Piccinelli, P. and Samuelsson, T. (2007) Evolution of the Iron-Responsive Element. *RNA*, **13**, 952-966. <http://dx.doi.org/10.1261/rna.464807>
- [19] Rogers, J.T., Bush, A.I., Cho, H.H., *et al.* (2008) Iron and the Translation of the Amyloid Precursor Protein (APP) and Ferritin mRNAs: Riboregulation against Neural Oxidative Damage in Alzheimer's Disease. *Biochemical Society Transactions*, **36**, 1282-1287. <http://dx.doi.org/10.1042/BST0361282>
- [20] Khan, M.A., Ma, J., Walden, W.E., *et al.* (2014) Rapid Kinetics of Iron Responsive Element (IRE) RNA/Iron Regulatory Protein 1 and IRE-RNA/eIF4F Complexes Respond Differently to Metal Ions. *Nucleic Acids Research*, **42**, 6567-6577. <http://dx.doi.org/10.1093/nar/gku248>
- [21] Haile, D.J., Rouault, T.A., Harford, J.B., *et al.* (1992) Cellular Regulation of the Iron-Responsive Element Binding Protein: Disassembly of the Cubane Iron-Sulfur Cluster Results in High-Affinity RNA Binding. *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 11735-11739. <http://dx.doi.org/10.1073/pnas.89.24.11735>
- [22] Beal, M.F. (2005) Mitochondria Take Center Stage in Aging and Neurodegeneration. *Annals of Neurology*, **58**, 495-505. <http://dx.doi.org/10.1002/ana.20624>

- [23] Basso, M., Samengo, G., Nardo, G., *et al.* (2009) Characterization of Detergent-Insoluble Proteins in ALS Indicates a Causal Link between Nitritative Stress and Aggregation in Pathogenesis. *PLoS ONE*, **4**, e8130. <http://dx.doi.org/10.1371/journal.pone.0008130>
- [24] Rogers, J.T., Mikkilineni, S., Cantuti-Castelvetri, I., *et al.* (2011) The Alpha-Synuclein 5'Untranslated Region Targeted Translation Blockers: Anti-Alpha Synuclein Efficacy of Cardiac Glycosides and Posiphen. *Journal of Neural Transmission*, **118**, 493-507. <http://dx.doi.org/10.1007/s00702-010-0513-5>
- [25] Zuker, M. (1989) Computer Prediction of RNA Structure. *Methods in Enzymology*, **180**, 262-288. [http://dx.doi.org/10.1016/0076-6879\(89\)80106-5](http://dx.doi.org/10.1016/0076-6879(89)80106-5)
- [26] Cho, H.H., Cahill, C.M., Vanderburg, C.R., *et al.* (2010) Selective Translational Control of the Alzheimer Amyloid Precursor Protein Transcript by Iron Regulatory Protein-1. *The Journal of Biological Chemistry*, **285**, 31217-31232. <http://dx.doi.org/10.1074/jbc.M110.149161>
- [27] Walden, W.E., Selezneva, A.I., Dupuy, J., *et al.* (2006) Structure of Dual Function Iron Regulatory Protein 1 Complexed with Ferritin IRE-RNA. *Science*, **314**, 1903-1908. <http://dx.doi.org/10.1126/science.1133116>
- [28] Qureshi, M., Brown, R.H., Rogers, I.T. and Cudkowicz, M.E. (2008) Serum Ferritin and Metal Levels as Risk Factors for Amyotrophic Lateral Sclerosis. *The Open Neurology Journal*, **2**, 51-54. <http://dx.doi.org/10.2174/1874205X00802010051>
- [29] Ferraiuolo, L., Kirby, J., Grierson, A.J., *et al.* (2011) Molecular Pathways of Motor Neuron Injury in Amyotrophic Lateral Sclerosis. *Nature Reviews Neurology*, **7**, 616-630. <http://dx.doi.org/10.1038/nrneurol.2011.152>
- [30] Rogers, J.T., Leiter, L.M., McPhee, J., *et al.* (1999) Translation of the Alzheimer Amyloid Precursor Protein mRNA Is Up-Regulated by Interleukin-1 through 5'-Untranslated Region Sequences. *The Journal of Biological Chemistry*, **274**, 6421-6431. <http://dx.doi.org/10.1074/jbc.274.10.6421>
- [31] Bandyopadhyay, S., Hartley, D.M., Cahill, C.M., *et al.* (2006) Interleukin-1 α Stimulates Non-Amyloidogenic Pathway by Alpha-Secretase (ADAM-10 and ADAM-17) Cleavage of APP in Human Astrocytic Cells Involving p38 MAP Kinase. *Journal of Neuroscience Research*, **84**, 106-118. <http://dx.doi.org/10.1002/jnr.20864>
- [32] Cahill, C.M., Lahiri, D.K., Huang, X. and Rogers, J.T. (2009) Amyloid Precursor Protein and Alpha Synuclein Translation, Implications for Iron and Inflammation in Neurodegenerative Diseases. *Biochimica et Biophysica Acta*, **1790**, 615-628. <http://dx.doi.org/10.1016/j.bbagen.2008.12.001>
- [33] Eisenstein, R., Tuazon, P.T., Schalinske, K.L., Anderson, S.A. and Traugh, J.A. (1993) Iron-Responsive Element-Binding Protein: Phosphorylation by Protein Kinase C. *The Journal of Biological Chemistry*, **268**, 27363-27370.
- [34] Zu, T., Gibbens, B., Doty, N.S., *et al.* (2011) Non-ATG-Initiated Translation Directed by Microsatellite Expansions. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 260-265. <http://dx.doi.org/10.1073/pnas.1013343108>
- [35] Mori, K., Weng, S.M., Arzberger, T., *et al.* (2013) The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTL/ALS. *Science*, **339**, 1335-1338. <http://dx.doi.org/10.1126/science.1232927>
- [36] Lee, Y.B., Chen, H.J., Peres, J.N., *et al.* (2013) Hexanucleotide Repeats in ALS/FTD form Length-Dependent RNA Foci, Sequester RNA Binding Proteins, and Are Neurotoxic. *Cell Reports*, **5**, 1178-1186. <http://dx.doi.org/10.1016/j.celrep.2013.10.049>
- [37] Xu, Z., Poidevin, M., Li, X., *et al.* (2013) Expanded GGGGCC Repeat RNA Associated with Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Causes Neurodegeneration. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 7778-7783. <http://dx.doi.org/10.1073/pnas.1219643110>
- [38] Shin, J., Charizanis, K. and Swanson, M.S. (2009) Pathogenic RNAs in Microsatellite Expansion Disease. *Neuroscience Letters*, **466**, 99-102. <http://dx.doi.org/10.1016/j.neulet.2009.07.079>
- [39] Fratta, P., Mizielińska, S., Nicoll, A.J., *et al.* (2012) C9orf72 Hexanucleotide Repeat Associated with Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Forms RNA G-Quadruplexes. *Scientific Reports*, **2**, 1016. <http://dx.doi.org/10.1038/srep01016>
- [40] Reddy, K., Zamiri, B., Stanley, S.Y., *et al.* (2013) The Disease-Associated r(GGGGCC) $_n$ Repeat from the C9orf72 Gene Forms Tract Length-Dependent Uni- and Multimolecular RNA G-Quadruplex Structures. *The Journal of Biological Chemistry*, **288**, 9860-9866. <http://dx.doi.org/10.1074/jbc.C113.452532>
- [41] Waite, A.J., Baumer, D., East, S., *et al.* (2014) Reduced C9orf72 Protein Levels in Frontal Cortex of Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration Brain with the C9ORF72 Hexanucleotide Repeat Expansion. *Neurobiology of Aging*, **35**, 1779 e5-1779 e13.
- [42] Bencze, K.Z., Kondapalli, K.C., Cook, J.D., *et al.* (2006) The Structure and Function of Frataxin. *Critical Reviews in Biochemistry and Molecular Biology*, **41**, 269-291. <http://dx.doi.org/10.1080/10409230600846058>
- [43] Chapdelaine, P., Coulombe, Z., Chikh, A., *et al.* (2013) A Potential New Therapeutic Approach for Friedreich Ataxia:

Induction of Frataxin Expression with TALE Proteins. *Molecular Therapy—Nucleic Acids*, **2**, e119.
<http://dx.doi.org/10.1038/mtna.2013.41>

- [44] Paige, J.S., Wu, K.Y. and Jaffrey, S.R. (2011) RNA Mimics of Green Fluorescent Protein. *Science*, **333**, 642-646.
<http://dx.doi.org/10.1126/science.1207339>
- [45] Yoshimura, S., Gerondopoulos, A., Linford, A., *et al.* (2010) Family-Wide Characterization of the DENN Domain Rab GDP-GTP Exchange Factors. *The Journal of Cell Biology*, **191**, 367-381. <http://dx.doi.org/10.1083/jcb.201008051>
- [46] Farg, M.A., Sundaramoorthy, V., Sultana, J.M., *et al.* (2014) C9ORF72, Implicated in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia, Regulates Endosomal Trafficking. *Human Molecular Genetics*, **23**, 3579-3595.
<http://dx.doi.org/10.1093/hmg/ddu068>