Neurochemistry International 100 (2016) 11–20

Contents lists available at ScienceDirect

Neurochemistry International

journal homepage: www.elsevier.com/locate/nci

Structural insight into C9orf72 hexanucleotide repeat expansions: Towards new therapeutic targets in FTD-ALS

Vijay Kumar [a], Tara kashav [b], Asimul Islam [a], Faizan Ahmad [a], Md. Imtaiyaz Hassan [a, *]

[a] Centre for Interdisciplinary Research in Basic Sciences, Jamia Millia Islamia, New Delhi, 110025, India
[b] Centre for Biological Sciences, Central University of South Bihar, Patna, 800014, India

ARTICLE INFO

Article history:
Received 10 June 2016
Received in revised form
12 August 2016
Accepted 12 August 2016
Available online 15 August 2016

Keywords:
C9orf72
G-quadruplexes
Amyotrophic lateral sclerosis
Frontotemporal dementia
C9-FTD/ALS
Therapy

ABSTRACT

Hexanucleotide repeat expansions, (G4C2) in the C9orf72 gene are considered as the single most common genetic cause of both frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS). (G4C2), either as DNA or the transcribed RNA, can folds into unusual secondary structures, including G-quadruplex, R-loop, I-motif and hairpin. These structural polymorphism at both DNA and RNA levels were proposed to initiate molecular cascade leading to ALS/FTD. G-quadruplexes are composed of stacked G4 tetrads, held by hydrophobic bonds, and is highly stable secondary structure. Here, we covers the structural and functional features of G-quadruplexes with an emphasis on C9orf72-repeat-associated FTD and ALS (C9-FTD/ALS). We also highlighted tools and techniques used to study the G-quadruplexes. Current perspectives for molecules that target G-quadruplexes as potential therapeutic are discussed. Our extensive analysis of structural features of G-quadruplexes will be used for a better understanding of molecular mechanism of C9-FTD/ALS.

Contents

1. Introduction

Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) are severe neurological diseases. ALS is common

* Corresponding author.
E-mail address: mihassan@jmi.ac.in (Md.I. Hassan).

http://dx.doi.org/10.1016/j.neuint.2016.08.008
0197-0186/

motor neuron disease initiated by the loss of motor neurons in the brain and spinal cord. It progresses to muscle weakness and paralysis and ultimately causes death due to respiratory failure (Kiernan et al., 2011; Kumar et al., 2016c; Rowland and Shneider, 2001). It occurs with an incidence of 2.6/100,000 per year (Nelson, 1995). FTD is the second most common cause of early onset dementia, leading to the degeneration of the frontal and temporal lobes of the brain. It occurs with an incidence of 3.5–4.1/100,000 per year in individuals under 65 (Harvey et al., 2003; Ratnavalli et al., 2002).

ALS and FTD are a continuum of two intersecting diseases, commonly mentioned as FTD/ALS (Kumar et al., 2016a). Firstly, motor dysfunctions have been noticed in 15% of FTD patients (Burrell et al., 2011; Lomen-Hoerth et al., 2002), and FTD is present in 15–18% of patients with ALS (Ringholz et al., 2005). Secondly, sporadic and familial cases of both disease are caused by an expansion of a hexanucleotide repeat of GGGGCC (HRE) in the first intron of the C9orf72 gene (DeJesus-Hernandez et al., 2011; Renton et al., 2011), termed C9-FTD/ALS. C9orf72 HRE is recognized as the most frequent genetic cause of both FTD and ALS as it accounts for up to 80% of familial FTD/ALS, 20%–50% of familial ALS, 5%–20% of sporadic ALS, and 10%–30% of FTD (Boeve et al., 2012; Chio et al., 2012; Cooper-Knock et al., 2012; Hsiung et al., 2012; Snowden et al., 2012). Healthy people carry an average of 2 repeats, with 90% having 8 or less (Rutherford et al., 2012). Among the ALS and FTD patients, the repeats range from 500 to several thousand units (Beck et al., 2013; van Blitterswijk et al., 2013). Clinically, patients with C9orf72 HRE are reported to have a higher incidence of bulbar-onset ALS, cognitive impairment with earlier disease onset, and accelerated progression compared with patients without the repeat expansion (Byrne et al., 2012; Millecamps et al., 2012; Stewart et al., 2012).

2. G-quadruplexes

The C9orf72 HRE DNA and RNA enable the formation of complex structures including G-quadruplexes (Fratta et al., 2012; Haeusler et al., 2014; Reddy et al., 2013). Guanosine-rich DNA and RNA sequences easily formed a stable four-stranded structure known as G-quadruplexes. Guanine bases are connected through Hoogsteen pairing to form a square planar or guanine tetrad structure. Two or more guanine tetrads subsequently stacked to form G-quadruplex which is stabilized by monovalent cations (Huppert, 2008) (Fig. 1). Both C9orf72 HRE DNA and RNA may contribute to the pathogenesis of ALS/FTD disease through a mechanism associated with their structure polymorphism (Vatovec et al., 2014).

2.1. Structural features of G-quadruplex

A G-quadruplex can be seen as an assembly of G-quartets, which is larger than Watson Crick base pairs of a double helix (Ou et al., 2008). G-quadruplexes display an extensive structural diversity and polymorphism as shown by X-ray crystallography and NMR spectroscopy. This polymorphism is generally depends on the composition of oligonucleotide sequence; number and orientation of oligonucleotide strands; type and size of loops; the angle of the glycosidic bonds; and the environment of the solution (Bidzinska et al., 2013). Some classification of G-quadruplexes is shown in Fig. 1.

2.2. Distribution of the G-quadruplex

Computational analysis of the human genome suggests that putative G-quadruplex forming sequences (PQS) contains a signature motif of G ≥ 3NxG ≥ 3NxG ≥ 3NxG ≥ 3Nx, with an estimate of ~376,000 motifs (Eddy and Maizels, 2006; Huppert and Balasubramanian, 2005; Kikin et al., 2006; Todd et al., 2005). These PQS are frequently present within human telomeric DNA sequences, ribosomal DNA sequences, transcription start sites (Du et al., 2008), the promoter (Huppert and Balasubramanian, 2007) and untranslated regions of mRNA (Bugaut and Balasubramanian, 2012), suggesting that G4 structures may control various cellular processes such as telomere maintenance, ribosome biogenesis, gene replication, transcription and translation. Recent advances in *in vivo* detection of G-quadruplex by structure-specific antibody and structure-binding ligands provide several convincing evidences for the existence of G-quadruplex structures at telomeres and in human chromosome (Lam et al., 2013).

3. Location and epidemiology of C9orf72 HREs

The C9orf72 gene is located on 9p21.2, and is formed by 11 exons, coding 3 transcription variants (V1, V2, V3) and two protein isoforms (a, b). Transcript V1 contains non-coding exon 1b, and V3 contains non-coding exon 1a, are fused to coding exons 2 through 5 and encode a longer isoform of the protein with 481 amino acid (called isoform a). Transcript V2 contains exon 1a which encodes a shorter protein with 222 amino acid (called isoform b) (Woollacott and Mead, 2014). The HRE lies in an intronic region of C9orf72 between non-coding exons 1a and 1b. V2 and V3 utilize exon 1a thus contains the HRE. However, V1 utilizes the alternative exon 1b thus exclude the HRE (Fig. 2).

The HREs observed at frequencies of up to 29% in FTD, 50% in ALS, and 88% in FTD/ALS patients (Cruts et al., 2013). The mutation frequencies were observed maximum in Caucasian populations of Europe and North America, and were markedly higher in ALS patients of Finland (Cruts et al., 2013; Majounie et al., 2012), Sweden (Smith et al., 2013), and Denmark (Lindquist et al., 2013). By contrast, the HREs were generally rare in Asian population (Cruts et al., 2013; Zou et al., 2013). In both FTD (AD and FTD Mutation Database) (Cruts et al., 2012) and ALS (ALS Online Genetics Database) (Abel et al., 2012), the frequency of HREs are considerably higher than that of any other disease-causing mutation. Besides Granulin (GRN) and microtubule-associated protein Tau gene (MAPT) which together explain 11–25% of familial FTD cases (DeJesus-Hernandez et al., 2011; Gijselinck et al., 2012), C9orf72 is third most frequently mutated FTD gene. In ALS, HREs occur 1–2-fold more often than SOD1 mutations (Smith et al., 2013; Stewart et al., 2012). The C9orf72 HREs were investigated in other neurodegenerative diseases such as Parkinson's, Alzheimer's, multiple sclerosis, schizophrenia, and other motor neuron diseases. However, they are rare and not the main cause of these disorders (Cacace et al., 2013; Cooper-Knock et al., 2013; Fenoglio et al., 2014; Galimberti et al., 2014; Harms et al., 2013; Hubers et al., 2014; Nuytemans et al., 2013).

4. Structural and functional features of C9orf72 HRE

G-quadruplex motifs are involved in two complex human diseases: cancers and neurological disorders. Physiological importance of G-quadruplexes is evident from these motifs which are distributed in human genome (Eddy and Maizels, 2006; Maizels and Gray, 2013), as well as their presence within microsatellites and regulatory elements in various genes (Huppert and Balasubramanian, 2005; Huppert et al., 2008; Todd et al., 2005). Two distinct mechanisms explain the involvement of G-quadruplexes in neurological diseases: (i) through repeat expansions of G-rich sequences that are predicted to form G-quadruplexes and are found to cause disease as seen in C9-FTD/ALS, and (ii) by mutations affecting the expression of G-quadruplex binding proteins