Seminars in Cell & Developmental Biology xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Seminars in Cell & Developmental Biology

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

Review

Regulation of a minimal transcriptome by repeat domain proteins

Nicola Ferreira, Oliver Rackham, Aleksandra Filipovska*

Harry Perkins Institute of Medical Research and The University of Western Australia, School of Molecular Sciences, Nedlands 6009, Perth, Western Australia, Australia

ARTICLE INFO

Article history:
Received 6 June 2017
Received in revised form 15 August 2017
Accepted 18 August 2017
Available online xxx

Keywords:
Pentatricopeptide repeat protein
Mitochondrial transcription termination factor
RNA-binding protein
Mitochondrial RNA
Gene expression

ABSTRACT

Repeat proteins regulate the expression of the mammalian mitochondrial genome at the level of transcription, processing, maturation, and translation. Defects in the regulation of mitochondrial gene expression due to mutations in genes encoding repeat proteins can lead to mitochondrial dysfunction and disease, however the molecular mechanisms that regulate mitochondrial gene expression and how defects in these processes cause disease still remains poorly understood. Recently solved crystal structures, characterisation of the new genetic models, and use of RNA sequencing (RNA-Seq) technologies have greatly expanded our current understanding of mitochondrial repeat proteins and biology.

Contents

1. Regulation of mammalian mitochondrial gene expression by modular proteins

The mammalian mitochondrial DNA (mtDNA) is a compact, circular, double-stranded genome, containing only 37 genes for 2 rRNAs, 22 tRNAs and 11 mRNAs. The mitochondrial mRNAs (mt-mRNAs) encode 13 polypeptides that are subunits of three respiratory complexes (Complex I, III and IV) and the ATP synthase, which together with Complex II form the oxidative phosphorylation (OXPHOS) system required for the majority of energy generation in cells (reviewed in [1,2]). The expression of mammalian mtDNA is regulated predominantly at a post-transcriptional level by many nuclear encoded RNA-binding proteins that are imported inside mitochondria post-translationally to transcribe, stabilise, modify, translate and degrade mt-RNAs [3]. The known roles of nuclear encoded mitochondrial RNA-binding proteins (mtRBPs) have been reviewed recently [1,2,4,5]. Here we discuss

* Corresponding author.
E-mail address: aleksandra.filipovska@uwa.edu.au (A. Filipovska).

http://dx.doi.org/10.1016/j.semcdb.2017.08.037
1084-9521/

the roles of two families of mtRBPs that consist of characteristic repetitive protein motifs, mitochondrial transcription termination factors (mTERF), that can bind mtDNA and mtRNA, and pentatricopeptide repeat (PPR) domain proteins, that bind mtRNA [6–8].

The MTERF proteins are a group of structurally similar proteins with variably repeated MTERF-motifs that form a half-donut-shape, right-handed superhelix, where the positively charged concave side forms a path for nucleic acid binding [9–12]. In humans there are four human MTERF proteins: MTERF1, MTERF2, MTERF3, and MTERF4, all of which reside inside mitochondria [6,9,13–17]. The molecular structures and roles of most mammalian MTERF proteins in different aspects of mitochondrial gene expression have been well characterized, although little is known about the role of MTERF2 in mitochondria besides its affinity for mtDNA binding and association with the nucleoid [6,9,13–18].

PPRs were initially discovered in plant organelles as a large family of up to 800 proteins [7,19] that were only recently identified to bind RNA in a sequence-specific and modular fashion [20–24]. In mammals there are only seven PPR proteins (Fig. 1) that are exclusively mitochondrially localized and regulate mitochondrial gene expression from transcription to protein synthesis: the mitochondrial RNA polymerase (POLRMT) [25], leucine-rich pentatricopeptide repeat cassette (LRPPRC) protein [26–29], mitochondrial RNase P protein 3 (MRPP3) [30–32], pentatricopeptide repeat domain proteins 1, 2, and 3 (PTCD1, PTCD2, and PTCD3) [33–36], and mitochondrial ribosomal protein of the small subunit 27 (MRPS27) [37].

MtDNA is transcribed by POLRMT as two long polycistronic transcripts that span almost the entire length of the heavy or light strands of the genome (Fig. 2) [38]. MTERF1 has been identified to bind within the $tRNA^{Leu(UUR)}$ gene [9,13,14,39] and *in vivo* studies have shown that its binding blocks transcription to prevent transcriptional interference at the light strand promoter (LSP) [15]. The two polycistronic transcripts are processed according to the tRNA punctuation model (Fig. 2), whereby mitochondrial rRNAs (mt-rRNAs) and mitochondrial mRNAs (mt-mRNAs) are interspersed by mitochondrial tRNAs (mt-tRNAs), which act as "punctuation marks" for processing [40,41]. Processing of the polycistronic transcripts involves the excision of mt-tRNAs at their 5′ ends by RNase P complex, which in mammals consists of three proteins, MRPP1, MRPP2, and MRPP3 [30–32,34,42], and at their 3′ ends by RNase Z, which is ELAC2 [34,43]. PTCD1 associates with ELAC2 to assist in mt-tRNA processing [34]. However, there are four exceptions to the tRNA punctuation model (Fig. 3) which include the 3′ end of *mt-ND6*, 5′ end of *mt-CO1*, 5′ end of *mt-CYB* and between *mt-ATP6* and *mt-CO3* [40]. It is not entirely clear which other enzymes process these non-canonical sites although PTCD2 has been suggested to be required for processing the 5′ end of the *mt-CYB* transcript. Loss of the fas-activated serine/threonine kinase 5 (FASTKD5) results in the accumulation of unprocessed transcripts from the 5′ end of *mt-CYB* and the 5′ end of *mt-CO1* and between *mt-ATP8/6-CO3* [35,44] and disruption of fas-activated serine/threonine kinase 4 (FASTKD4) has also been shown to cause the accumulation of the *mtND5-CYB* precursor transcript. It has been suggested that the FASTKD protein family contain a putative endonuclease in the RNA-binding domain, RAP [45]. However, it still remains to be confirmed if PTCD2 has endonuclease activity or recruits endonucleases that are not yet characterized or identified.

The cleaved RNA transcripts then undergo maturation mediated by the LRPPRC/SLIRP complex [28,46] which includes polyadenylation of mt-mRNAs by the poly(A) polymerase (PAP), modification of tRNAs by a range of enzymes [4] and methylation of rRNAs by a variety of enzymes, including NSUN4 [6,11,47], followed by translation on mitochondrial ribosomes [48–50]. Mitochondrial ribosomes contain two PPR proteins within the small ribosomal subunit (PTCD3 and MRPS27) and are assembled co-transcriptionally [42]. Their assembly is mediated by a number of different factors, some characterized such as DDX28 [51] and MTERF3 [17], with additional players yet to be characterized.

The variation in the levels of mt-RNAs underlies the significance of RNA-binding proteins in the post-transcriptional regulation of mitochondrial gene expression [52]. Repeat domain proteins play key roles in every aspect of mitochondrial gene expression, from transcription to protein synthesis, and mutations or loss of these proteins can lead to mitochondrial dysfunction and disease [6,15,16,26,28,42,53,54]. Here we review the current advances towards understanding the role of mammalian MTERF and PPR proteins in mitochondrial biology and pathology.

2. Mitochondrial transcription termination factors (MTERFs)

2.1. Mitochondrial transcription termination factor 1 (MTERF1)

MTERF1 contains 8 MTERF-motifs, comprising of 19 α-helices [9]. MTERF1 binds in a sequence-specific manner to a 28 bp region within the $tRNA^{Leu(UUR)}$ gene, downstream of the 16S mt-rRNA region and was originally thought to promote heavy-strand transcription termination *in vitro* [13]. However, the role of MTERF1 in heavy-strand transcription termination has since been debated as MTERF1 has been shown to partially terminate transcription from the heavy strand promoter (HSP), while completely terminating transcription from the LSP [14]. The role of MTERF1 in light-stand transcription termination was identified *in vivo* when both copies of the mouse *Mterf1* genes were deleted to reveal that MTERF1 is dispensable for rRNA gene transcription regulation and instead it is required, but not essential, for minimizing transcription interference at the light strand promoter [15]. A mechanism of transcription termination by MTERF1 has since been proposed where MTERF1 binds to the major groove of DNA at its target site in the $tRNA^{Leu(UUR)}$ gene to promote unwinding of the DNA helix and eversion of several key residues to cause base flipping in a stepwise manner, which is critical for transcription termination [9,55]. In addition to transcription termination, MTERF1 was proposed to promote rRNA synthesis by binding to both the heavy strand transcription initiation and termination sites to mediate the formation of a DNA loop [39]. A recent report has also implicated MTERF1 in the regulation of mtDNA replication pausing *in vitro* [56,57], indicating that overall MTERF1 plays an important role in regulating mitochondrial transcription and mtDNA replication.

MTERF1 binds DNA in a site and cell specific manner [9,52], with conserved residues responsible for the sequence-specific binding [9]. Pathogenic mutations in these residues in MTERF1 have been proposed to interfere with MTERF1 binding and transcriptional termination [9,58]. Furthermore, MTERF1 has also been shown to have a tissue-specific transcriptional and translational compensatory role in response to decreased mtDNA copy number in a *MPV17* knockout mouse model that causes severe mitochondrial DNA depletion [59]. This implicates MTERF1 in diseases caused by impaired mitochondrial gene regulation, and further studies will shed light on its tissue-specific role in disease.

2.2. Mitochondrial transcription termination factor 3 (MTERF3)

MTERF3 is a highly conserved protein that contains 7 MTERF-motifs, consisting of 22 α-helices [10,60]. MTERF3 has a concave side with a strong positive charge that represents a nucleic-acid binding site, important for its role in binding mtDNA to regulate mitochondrial gene expression [10]. MTERF3 is essential for life and heart-specific knockout of mouse *Mterf3* gene causes respiratory chain defects and consequently mitochondrial dysfunction which

Please cite this article in press as: N. Ferreira, et al., Regulation of a minimal transcriptome by repeat domain proteins, Semin Cell Dev Biol (2017), http://dx.doi.org/10.1016/j.semcdb.2017.08.037