Fungal Genetics and Biology 114 (2018) 34–41

Contents lists available at ScienceDirect

Fungal Genetics and Biology

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

Comparative genomic analysis of fungal TPP-riboswitches

Mikhail A. Moldovan[a,b,c,*], Svetlana A. Petrova[a], Mikhail S. Gelfand[a,b,c,d]

[a] A.A. Kharkevich Institute for Information Transmission Problems, RAS, Bolshoy Karetny per. 19, Moscow 127051, Russia
[b] Faculty of Bioengineering and Bioinformatics, M.V. Lomonosov Moscow State University, Vorobievy Gory 1-73, Moscow 119991, Russia,
[c] Center for Data-Intensive Biomedicine and Biotechnology, Skolkovo Institute of Science and Technology, Moscow 143028, Russia
[d] Faculty of Computer Science, Higher School of Economics, Kochnovsky pr. 3, Moscow 125319, Russia

ARTICLE INFO

Keywords:
Riboswitch
Fungi
Thiamin
TPP
Regulation
Alternative
Splicing

ABSTRACT

Riboswitches are conserved RNA structures located in non-coding regions of mRNA and able to bind small molecules (e.g. metabolites) changing conformation upon binding. This feature enables them to function as regulators of gene expression. The thiamin pyrophosphate (TPP) riboswitch is the only type of riboswitches found not only in bacteria, but also in eukaryotes – in plants, green algae, protists, and fungi. Two main mechanisms of fungal TPP riboswitch action, involving alternative splicing, have been established so far. Here, we report a large-scale bioinformatic study of riboswitch structural features, action mechanisms, and distribution along the fungal taxonomy groups. For each putatively regulated gene, we reconstruct the riboswitch structure, identify other components of the regulation machinery, and establish mechanisms of riboswitch-mediated regulation. In addition to three genes known to be regulated by TPP riboswitches, thiazole synthase THI4, hydroxymethilpyrimidine-syntase NMT1, and putative transporter NCU01977, we identify two new genes, a putative thiamin transporter THI9 and a transporter of unknown specificity. While the riboswitch sequence and structure remain highly conserved in all species and genes, the mode of riboswitch-mediated regulation varies between regulated genes. The riboswitch usage varies strongly between fungal taxa, with the largest number of riboswitch-regulated genes found in Pezizomycotina and no riboswitch-mediated regulation established in Saccaromycotina.

1. Introduction

Riboswitches, RNA structures capable of changing the conformation in response to binding by small molecules and using this mechanism to regulate gene expression have been initially identified in bacteria. While it had been shown that biosynthetic genes of vitamins B_1 (thiamin), B_2 (riboflavin), and B_{12} (cobalamin) are inhibited when the respective metabolite is present in the medium (Kreneva et al., 1997; Miranda-Rios et al., 2001; Nou and Kadner, 1998; Ravnum and Andersson, 2001), no candidate protein transcription factors could be found. Eventually, the mRNA sequence analysis yielded conserved mRNA sequences with highly conserved secondary structures, and it had been proposed, and then experimentally validated, that these structures can directly and specifically bind metabolites, change conformation and thus control gene expression (Gelfand et al., 1999; Miranda-Rios et al., 2001; Nahvi et al., 2002; Rodionov et al., 2002; Vitreschak et al., 2002; Winkler et al., 2002).

Other types of riboswitches identified in early studies using sequence analysis are S-boxes (now known as SAM-riboswitches) initially observed in Gram-positive bacteria as highly conserved structures upstream of genes involved in the methionine and cysteine biosynthesis and then shown to cause transcription termination upon binding of *S*-adenosylmethionine (Epshtein et al., 2003; Grundy and Henkin, 1998; McDaniel et al., 2003; Winkler et al., 2003), LYS-elements, regulating biosynthesis of lysine (Rodionov et al., 2003), and guanine riboswitches regulating guanine biosynthesis (Batey et al., 2004).

Over the years that followed the discovery of riboswitches and the basic principles of their action, diverse and intricate mechanisms of riboswitch-mediated regulation in bacteria have been established (Sudarsan et al., 2006; Winkler et al., 2004) and new types of riboswitches as well as their ligands have been found (Baker et al., 2012; Cromie et al., 2006; Dann et al., 2007; McCown et al., 2017). Based on the simplicity of the regulatory mechanism not requiring additional components and wide taxonomic distribution, Nahvi et al. (2002) and Vitreschak et al. (2004) suggested that riboswitches could be the earliest gene regulatory mechanism.

As TPP is the end product of thiamin (vitamin B1) metabolism, branches of the thiamin metabolism are regulated by TPP riboswitches.

* Corresponding author at: A.A. Kharkevich Institute for Information Transmission Problems, RAS, Bolshoy Karetny per. 19, Moscow 127051, Russia.
E-mail address: mika.moldovan@gmail.com (M.A. Moldovan).

https://doi.org/10.1016/j.fgb.2018.03.004
Received 8 October 2017; Received in revised form 17 February 2018; Accepted 8 March 2018
Available online 13 March 2018

Bacterial TPP riboswitches regulate either transcription or translation via TPP-dependent assembly of terminator or sequestor hairpins, respectively. In archaea (*Thermoplasma* spp.) and in some bacteria (in particular, actinobacteria and cyanobacteria) TPP-riboswitch regulation does not seem to involve formation of additional structures, as these riboswitches directly hide gene starts (Rodionov et al., 2002).

Shortly after the initial discovery of prokaryotic riboswitches it has been shown that fungus *Aspergillus oryzae* has two highly conserved motifs in the 5′-UTR of the *THIA (THI4)* gene encoding a thiamin biosynthesis enzyme (Kubodera et al., 2003). *THI4* is activated at low thiamin concentrations and repressed when thiamin is abundant. Deletion of either of these regions resulted in compromised splicing, which is essential for the *THI4* expression. Structural alignment of *THI4* 5′-UTR regions from fungal genomes yielded a riboswitch-like structure (Kubodera et al., 2003). Later, TPP riboswitches were found also in green algae (Croft et al., 2007) and higher plants (Bocobza et al., 2007).

Apparently, uncoupling of transcription and translation in eukaryotes and specific mechanisms of the termination of transcription and initiation of translation, as well as emergence of the splicing stage in gene expression resulted in changes in the mechanisms of riboswitch regulation. Whereas bacterial riboswitches mostly operate at the stages of transcription or translation, eukaryotic riboswitches modulate alternative splicing (in fungi and green algae) or influence mRNA stability (in higher plants) (Bocobza et al., 2007; Croft et al., 2007; Sudarsan et al., 2003).

Two mechanisms of fungal riboswitch action have been identified and verified experimentally in *Neurospora crassa* (Cheah et al., 2007; Li and Breaker, 2013). In the case of the *NMT1* riboswitch, at low TPP concentrations (Fig. 1a), complementary base pairing between the P4–P5 riboswitch region and the region containing the downstream splice site results in the use of the upstream splice site to excise the larger intron containing an upstream ORF (uORF). This enables *NMT1* ORF to be translated. When the TPP concentration is high (Fig. 1b), the riboswitch forms a stable complex with a TPP molecule making the downstream donor splice site available for splicing, which results in partial suppression of the splicing branch site and excision of the smaller intron. This allows the short uORF to be translated, and subsequently the *NMT1* translation is suppressed (Cheah et al., 2007).

In the case of putative transporter *NCU01977* (Fig. 2), the riboswitch is formed in the first intron, which has two alternative donor splice sites. The region α localized between these sites is complementary to the downstream part of the basal (P1) riboswitch stem (region α′). At low TPP concentration, region α interacts with region α′, leading to excision of the larger intron and active translation. At high TPP concentrations, the riboswitch adopts a stable ligand-bound conformation and the absence of the α–α′ interaction releases the

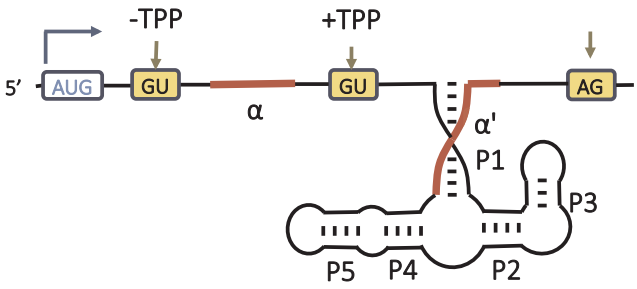

Fig. 2. Structures in the first intron of *NCU01977* involved in the regulation its expression in *Neurospora crassa*. Complementary regions α and α′ interacting at low TPP concentrations are shown as thicker lines. This interaction leads to excision of the larger intron, utilizing the upstream donor splice site. At high TPP concentrations, splicing mainly occurs from the downstream donor splice site. Modified from Li and Breaker (2013). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

downstream donor splice site; hence, the shorter intron is excised, that leads to translation of an aberrant peptide (Li and Breaker, 2013).

A large number of available, completely sequenced fungal genomes allowed us to perform a large-scale comparative-genomics analysis to identify sets of genes regulated via TPP-riboswitches in diverse taxa, characterize the mechanisms of riboswitch-mediated regulation for different genes, describe conserved features of TPP-riboswitches and other relevant structures and sequence elements.

2. Materials and methods

The TPP-riboswitch structural pattern was obtained from Rfam (Griffiths-Jones et al., 2003). InfeRNAl cmsearch (Nawrocki et al., 2009) was used for pattern search. Hits were refined manually and using the Zuker algorithm, integrated in the program Mfold (Zuker, 2003), and InfeRNAl cmalign (Nawrocki et al., 2009). Fungal genomes and annotations were obtained from the NCBI database (Benson et al., 1998). AUGUSTUS (Stanke and Waack, 2003) was used for automatic annotation of unannotated genomes. InfeRNAl cmsearch was used to find candidate riboswitches in all genomes. Then genes associated with putative riboswitches were extracted and clustered into orthology groups using BLAST-all (Altschul et al., 1990). BLAST was used to add unregulated homologs to these orthology groups.

Functional annotation was done using BLASTp and TCDB (Saier et al., 2016). Transmembrane segments in putative transporters were identified with TMHMM (Krogh et al., 2001). Transcripts and protein sequences were obtained from the Ensembl fungi database (Kersey et al., 2010). We refined our annotations of regions adjacent to putative

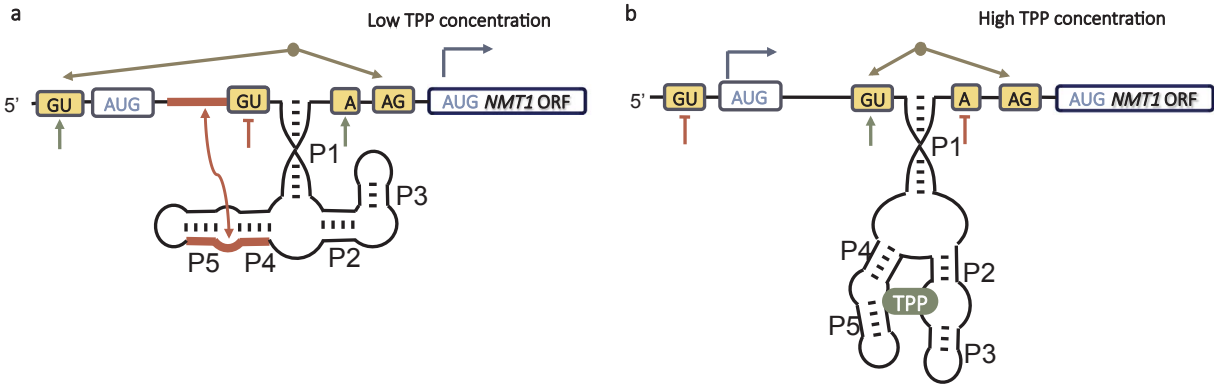

Fig. 1. Mechanism of the *NMT1* riboswitch action in *Neurospora crassa*. At low TPP concentrations (a), the P4–P5 riboswitch region interacts with the region containing the downstream donor splice site, suppressing it. Complementary regions are shown as thicker lines, complementary interaction is shown with the double arrow. Suppression is indicated with the blunted arrow, activation, with an arrow below the corresponding site. Translation start is marked with the arrow above the AUG codon. At high TPP concentrations (b), the upstream donor splice site is inactive and the splicing branch site is suppressed, leading to the mRNA with translated uORF and suppressed translation of the main ORF. Modified from Cheah et al. (2007). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)