

How to Have Dual Lives: Proteins that Bind DNA and RNA

*Nora Beatriz Calcaterra

Institute of Molecular and Cellular Biology of Rosario (IBR), National Council for Scientific and Technical Research (CONICET) - Faculty of Biochemical and Pharmaceutical Sciences, National University of Rosario (UNR) and Esmeralda Ocampo (S2000FHQ) Rosario, Argentina

***Corresponding author:** Nora Beatriz Calcaterra, Institute of Molecular and Cellular Biology of Rosario (IBR), National Council for Scientific and Technical Research (CONICET) - Faculty of Biochemical and Pharmaceutical Sciences, National University of Rosario (UNR) and Esmeralda Ocampo (S2000FHQ) Rosario, Argentina. Tel: 341-435-0661; E-mail: calcaterra@ibr-conicet.gov.ar

Received date: Jul 25, 2014; **Accepted date:** Jul 26, 2014; **Published date:** Jul 27, 2014

Copyright: © Calcaterra NB. This is an open-access editorial distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Editorial

Gene expression comprises transcription, splicing, mRNA export, mRNA stability and translation, which come together to control the abundance of any protein. These steps are not independent but are linked by proteins that act on multiple levels of gene regulation by interacting with the nucleic acids.

Traditionally, Nucleic Acid Binding Proteins (NABPs) were classified in several types according to the nature of the nucleic acid they bound, as well as to their sequence specificity. Thereby, it is frequent to find NABPs classified as DNA-binding proteins, DNA-binding proteins that specifically bind single-stranded DNA, RNA-binding proteins that bind to single- or double-stranded RNA; as well as proteins that bind specific or non-specifically to nucleic acids. On the other hand, proteins binding to DNA or RNA have been often considered and studied as separate functional classes of NABPs since Transcription Factors (TF) bind to genomic promoters and control target gene expression by activating or repressing RNA polymerase activity, and RNA binding proteins modulate protein expression by regulating the degradation, stability, and translation of target mRNAs. However, the consideration of DNA- and RNA-binding functions within proteins as discrete units has been rapidly changing in recent times. Several articles have demonstrated the dual DNA- and RNA-binding capability and cytoplasm-nucleus location of a growing group of NABPs. This duality creates an important additional level of gene regulation that serves to connect transcriptional and post-transcriptional control, and suggests a key role of these proteins in modulating gene expression, survival, and cellular homeostasis.

Dual NABPs may participate in gene expression by acting in different ways. TFs with dual binding abilities can establish an autoregulatory feedback loop in gene expression. TF with similar affinity for DNA and RNA may be kidnapped by nascent RNA molecules, thus reducing promoter occupancy and transcription of target genes. In addition to binding to the promoter of target genes, several NABPs may participate in several steps of the RNA metabolism (miRNA processing, mRNA stability, degradation, and translation). Further, some dual NABPs can bind and modulate secondary structures adopted by DNA and RNA during transcription and/or RNA processing, thus promoting or impairing transcription/translation processes.

The typical TF responsible for autoregulatory loop in gene expression regulation is TFIID. TFIID is perhaps the first protein identified as having dual nucleic acid binding capability. It was initially isolated from a 7S RNP complex as bound to 5S rRNA in *Xenopus laevis* oocyte cell extracts [1]. Later, it was also found as a transcription factor capable of activating 5S rRNA gene expression [2,3]. TFIID escorts 5S

rRNA during nuclear export as 7S RNP complex and remains in the cytoplasm until the 5S rRNA is released for ribosome assembly. Release of the 5S rRNA allows TFIID to return to the nucleus, where it can activate the expression of additional 5S rRNA. As 5S rRNA transcripts are produced, TFIID progressively becomes sequestered in the cytoplasm thus leading to reduced transcription of the 5S rRNA gene.

TF also regulate genes post-transcriptionally through the regulation of miRNA biogenesis. Several TF have been shown to regulate pri-miRNA processing by Drosha, a key step in the biogenesis of functional miRNAs. Often, identical sequences in DNA and RNA recruit NABP for transcriptional and posttranscriptional gene regulation.

During transcription, double-stranded DNA is open exposing single-stranded regions that can adopt secondary structures that may block or favour transcription initiation. In the past few decades a large number of scientific papers, both computational and experimental in vitro and in vivo; have demonstrated that G-rich DNA sequences can fold as intra-strand quadruple helix structures called G-quadruplexes or G4-DNA. G4-structures can be also formed in RNA molecules, usually in the 5'UTR. Several NABPs have been identified binding to G4-DNA and G4-RNA of the same gene and modulating their foldings. The presence of G4 might serve as a mark for recruiting specific TF or other RNA-binding proteins, thus promoting or avoiding the progress of transcription or translation as well.

Elucidating the mechanisms of action of these multifunctional proteins is difficult and could explain the slow progress in the field to date. The identification of the pool of dual NABPs in eukaryotic cells is not easy. For human cells, the QuickGO gene ontology database (European Bioinformatics Institute) yields low amount of proteins having experimental evidences for both RNA binding and DNA binding. Despite the PROTEOME database (BioBase) returns a higher amount of these proteins, direct experimental evidences regarding this dual activity is lacking for many of these entries. Recently, it was reported the "mRNA interactome" of HeLa cells consisting of 860 mRNA-binding proteins [4]. Among them, approximately one half had been previously characterized as double-stranded DNA binding proteins [5], thus suggesting a relatively high representation in the human proteome. The number of dual NABPs would likely increase when including proteins present in other cell types. Certainty, the advent of new technologies ensures that more dual-NABPs and their mediated functions will be discovered in the coming years. This could be a focus of future research to enable a better understanding of how these highly abundant proteins fit into the broader theme of gene regulation.

References

1. Picard B, Wegnez M (1979). Isolation of a 7S particle from *Xenopus laevis* oocytes: a 5S RNA-protein complex. *Proc Natl Acad Sci U S A*. 76: 241-5.
2. Engelke DR, Ng SY, Shastry BS, Roeder RG (1980). Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. *Cell*. 19: 717-28.
3. Pelham HR, Brown DD (1980). A specific transcription factor that can bind either the 5S RNA gene or 5S RNA. *Proc Natl Acad Sci U S A*. 77: 4170-4.
4. Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell*. 149: 1393-406.
5. Hu S, Xie Z, Onishi A, Yu X, Jiang L (2009). Profiling the human protein-DNA interactome reveals ERK2 as a transcriptional repressor of interferon signaling. *Cell*. 139: 610-22.