

Catalytic RNA world relics in Dicer RNAs

Sayak Ganguli*, Dey S.K., Priyanka Dhar, Protip Basu, Paushali Roy and Abhijit Datta

DBT-Centre for Bioinformatics, Presidency College, Kolkata, sayakbif@yahoo.com

Abstract- RNA interference (RNAi) is a naturally occurring phenomenon of RNA-mediated gene silencing that is highly conserved among multicellular organisms. In the first step of the pathway, long double-stranded RNA molecules are chopped into shorter duplexes with 2 nucleotide overhangs at both 3' ends by an endonuclease dubbed Dicer, the structure of which has been solved only recently. This results in the formation of small 21 nucleotide long RNAs, aptly named small or short interfering RNAs (siRNAs), which are incorporated into a multimeric protein complex, the RNA-induced silencing complex (RISC). One of the two-siRNA strands guides RISC to a complementary RNA. After hybridization the endonucleolytic "slicer" activity of RISC cleaves the target RNA, thus preventing its translation. While long double-stranded RNA molecules can be employed to induce RNAi in lower eukaryotes, siRNAs being 21 nucleotides in length have to be used for gene silencing in mammalian cells in order to prevent the activation of an unspecific interferon response [1]. In contrast to siRNAs, however, miRNAs are capable of inhibiting translation of the targeted mRNA without degrading it (at least in mammalian cells)[2-4]. The need for in silico analysis of the components of the RNA interference pathway arises from the fact that very little is known about the structural and interacting properties of the components. With the above background the analysis was performed to identify putative catalytic motifs in the mRNA of the DICER enzyme.

Key Words: RNAi, Drosha, RISC, miRNA, ribozymes, motifs

Introduction

Over the past two decades it has become clear that a variety of the RNA molecules have important or essential biological functions in cells. RNA is proficient at forming complex and varied tertiary structures as revealed by high resolution structures of a handful of RNAs. The secondary and tertiary structures of RNA are key for understanding their biological activity. Motifs that stabilize RNA 3D folds are relatively small and often involve backbone functional groups, making them impossible to detect even when in large families of secondary structures. Tetraloops and their receptors, U – turns, dinucleotide platforms, ribose zippers and S – turns all consist of 4 – 11 nucleotides and occur within a variety of sequence contexts. In addition, non – canonical base pairs often create context dependent helical geometries or surfaces used in RNA – RNA and RNA – Protein recognition. One of the most essential structures of RNA is the RNA hairpin. It can guide RNA folding, determine interactions in a ribozyme, protect messenger RNA from degradation, serve as recognition motif for RNA binding proteins and act as substrates for enzymatic reactions. Eukaryotic small RNAs of approximately 21–24 nucleotides function as guide molecules in a remarkably wide range of biological processes, including developmental timing and patterning, formation of heterochromatin, genome rearrangement, and antiviral defense [6-7]. They belong to at least two general classes, miRNA and siRNA. miRNAs (approximately 21–22 nucleotides) are found in plants and animals and are often phylogenically conserved within their respective kingdoms. These miRNAs are formed from a precursor which is transcribed from genes which are non protein coding. A part of this nascent precursor

adopts a fold-back structure that interacts with a multidomain RNaseIII-like enzyme termed DICER or DICER-LIKE (DCL1 in Arabidopsis), which catalyzes accurate excision of the mature miRNA[8]. The miRNAs then associate with ribonucleoprotein complexes that function to negatively regulate target genes controlling a range of developmental events, such as timing of cell fate decisions, stem cell maintenance, apoptosis, organ morphogenesis and identity, and polarity [6]. siRNAs are chemically similar to miRNAs, although in plants they typically range in size between 21 and 24 nucleotides [9-11]. They are associated with both post-transcriptional forms of RNA interference and transcriptional silencing involving chromatin modification [7]. siRNAs are processed from precursors containing extensive or exclusive double-stranded RNA (dsRNA) structure, such as transcripts containing inverted repeats or intermediates formed during RNA virus replication. siRNA precursors can also be formed by the activity of one or more cellular RNA-dependent RNA polymerases (RdRp), as was shown genetically in several screens for RNA silencing-defective mutants [13-16]. Arabidopsis plants contain at least three active RdRp genes, termed RDR1, RDR2, and RDR6 (also known as SDE1/SGS2) [14,15 and 19]. RDR6 is necessary for sense transgene mediated RNAi, but not for silencing of constructs that encode transcripts with hairpins containing extensive dsRNA structure [14,15 and 19]. In many animals, both miRNAs and siRNAs are formed by the activity of the same DICER enzyme [20-27], although in plants they are formed by distinct DCL activities [7]. Arabidopsis contains four DCL genes (DCL1 to DCL4), only one of which (DCL1) has been assigned a