Ménage à Trois

Intimate Relationship Among a MicroRNA, Long Noncoding RNA, and mRNA

Gerald W. Dorn II, Scot J. Matkovich

MicroRNAs are one of multiple species of noncoding RNAs that also include lncRNAs. Because RNAs do not possess intrinsic catalytic activity, biological function is directly or indirectly determined by nucleotide sequence. For example, 20-nucleotide linear microRNAs embedded with Argonaute proteins in RNA-induced silencing complexes bind via seed sequence interactions to complementary nucleotides in 3′ untranslated regions (UTRs) of protein-coding mRNAs, thereby recruiting them to the RNA-induced silencing complex for silencing and degradation. Primary microRNA nucleotide sequence is therefore a key determinant of mRNA targeting.

Most bioinformatic miR–mRNA sequence comparison algorithms, such as TargetScan (www.targetscan.org), do not consider the physical configurations resulting from pairing of internal complementary mRNA sequences. Obviously, in addition to having complementary sequence, a microRNA-binding domain must be structured such that the critical seed sequence–binding site is accessible, that is, at least partially single stranded. This is depicted in Figure 1 for miR-499 and its validated mRNA target, the transcription factor Sox6.7,8 TargetScan 6 identifies 4 potential miR-499–binding sites within the mouse Sox6 3′ UTRs, one of which is poorly conserved and thermodynamically unfavorable and therefore not depicted. Structural modeling of the Sox6 3′ UTRs using MFold (mfold.rit.albany.edu) reveals that seed region–binding sites of 2 of the 3 remaining binding domains are more accessible (ie, have a greater open probability, $P_o=58\%$ and 41%) than the other ($P_o=24\%$; Figure 1). mRNA structure is therefore another important determinant of miR–mRNA binding.

Any RNA–RNA pairing event must follow the law of mass action, driven by the amounts of free microRNA and mRNA, and the RNA–RNA binding energy (analogous to binding affinity; lower binding energy reflects greater binding affinity).9,10 Thus, quantity of microRNA and mRNA and the thermodynamic characteristics of their pairing are also key determinants of miR–mRNA binding. RNA mass is measurable by deep sequencing,11–13 and raw and aligned microRNA and mRNA mouse heart sequence data are publicly available for interrogation (www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE55792). This resource reports miR-499 levels as $\approx 1600$ reads/million miR reads and Sox6 mRNA levels as $\approx 3$ to 6 fragments/reads per kilobase of exon per million mRNA reads (FPKM), an expression level typical for transcription factors. The minimum free energy of hybridization for miR-499 to the Sox6 3′ UTRs (calculated using RNAhybrid [bibiserv.techfak.uni-bielefeld.de/rnahybrid])14 is $\approx 21$ kcal/mol for each site, reflecting strong binding (Figure 1). Thus, miR-499 and Sox6 are both present at
meaningful levels in mouse hearts, the structure of the mouse Sox6 3′ UTRs provides a high probability of access to ≥2 miR-499–binding sites, and miR–mRNA binding energy at these sites is compatible with formation of an RNA pair with sufficient stability to retain the mRNA at the microRNA-loaded RNA-induced silencing complex. Accordingly, Sox6 is highly (≈5-fold) enriched in cardiac RNA-induced silencing complexes.13,15

The same analytic tools can be applied to the lncRNA–microRNA–mRNA interaction described by Wang et al.3 TargetScan identifies Myd88-binding sites for ≈80 microRNAs, but not for miR-489 (likely because of poor seed sequence complementarity). RNAhybrid analysis of the miR-489 site described by Wang et al3 reveals a healthy minimum free energy of hybridization between miR-489 and Myd88 of −19.2 kcal/mol. However, structural modeling of the Myd88 mRNA 3′ UTRs shows that the reported miR-489–binding site is relatively inaccessible (P = only 16%) and that the open seed binding configuration is still relatively closed (Figure 2, right).

High mass can compensate for low binding affinity. Thus, an improbable binding event can have physiological significance if the 2 binding partners are present at large quantities. According to the same web resource used above, Myd88 mRNA is expressed at levels similar to Sox6 mRNA (Figure 2). However, mouse heart miR-489 levels were not measurable (ie, <1 read/10 million miR reads) in published reports,11,12 and the miR-489 host gene (Calcr) is also present at <1 read/10 million. Thus, miR-489 is not common in normal mouse hearts. Given the unfavorable mRNA 3′ UTR structure, its interaction with native Myd88 may also be uncommon.

We applied the same tools and resources to the miR-489–lncRNA AK048451 interaction (Figure 2, left). AK048451 is present in (poly-A selected) normal adult mouse heart mRNA at <1 read/10 million mRNA reads (although we consistently detect it at levels of 2–3 reads/10 million mapped reads in embryonic mouse hearts; Dorn and Matkovich, unpublished data, 2014). Accessibility of the seed sequence–binding site at the reported miR-489–binding domain in lncRNA AK048451 is comparable with the functional miR-499 sites in Sox6 (P = 35%), with an equivalent minimum free binding energy (−21.3 kcal/mol). When this lncRNA is upregulated, as in angiotensin II–stimulated hypertrophy, it is likely to bind to miR-489.

Limitations and assumptions of these analytic tools are significant: we assume that the measured level of each RNA species represents its free concentration in the cell of phenotypic interest (cardiomyocyte) and that the microRNA, mRNA, and lncRNA are not subject to subcellular compartmentalization, that is, are mutually available to interact. The model also does not account for sequestration binding of miR-489 by any other lncRNAs or more abundant cardiac-expressed mRNAs with roughly equivalent binding energies for miR-489 (eg, Mef2c: FPKM ≈25; binding energy for miR-489=−17.1 kcal/mol).

These analyses illustrate how important variables can be overlooked when an inherently rich and necessarily complex biologically regulatory mechanism is reduced to a linear concept. Not only will multiple variables not be accounted for, but each of the variables is context dependent, transient, and probabilistic. Thus, RNA–RNA interactions, whether intramolecular binding that produces folded RNA structures or...
intermolecular binding between microRNAs and their targets or sponges, are dynamic. Accordingly, RNA structures are ranked nondeterministically using either the open probability metric or by calculating the RNA–RNA minimum free energy of binding. Furthermore, the steady-state levels of RNA species, whether coding and noncoding, are constantly changing because of modulated expression, processing, and sequestration. Absolute RNA content (mass), not relative RNA expression, is a key factor. These quantitative data are not provided by arrays or real-time quantitative polymerase chain reaction. Finally, individual noncoding RNAs have multiple different (DNA and RNA) binding partners and therefore exhibit multifunctionality. Indeed, the complex folded structures of lncRNAs produce striking mechanistic diversity, enabling them to function as protein anchors, as chaperones that target transcriptional modulators to specific genes, and as binding partners for other RNA species.

As shown above, Web-based resources can predict RNA structure and thermodynamic parameters of RNA–RNA binding, and mRNA and small RNA sequencing data are increasingly obtainable. The National Center for Biotechnology Information GEO resource to which we uploaded results of >40 individual normal adult mouse heart RNA and small RNA sequencing studies provides access to preprocessed data (facilitating interrogation of specific transcripts of interest) and raw sequencing reads (for application of improved or alternate alignment algorithms).

With time, we will discover additional mechanisms that modify gene expression and transcript stability/function. A lncRNA acting as a microRNA sponge is the current example of one epigenetic mechanism regulating another, so-called interepigenetic regulation. Indeed, by directing proteins to specific genomic DNA sites, lncRNAs are well positioned to modify other epigenetic events, such as chromatin structure and DNA methylation. We propose that regulatory complexity be embraced: The observation that a given molecular or functional interaction can occur should be supported by analyses of the probability that it takes place in the relevant biological context. Unbiased genome-wide experimentation that quantitatively evaluates levels of, and interactions between, microRNAs, IncRNAs, and mRNAs is a first step toward this goal. Agnostic results ranked according to probability estimates factoring in RNA structure, binding affinity, and RNA species abundance can be validated using standard coprecipitation and function reporter assays with binding site mutagenesis. Such an approach has the same advantages as does eschewing candidate gene analysis in favor of unbiased genome-wide discovery of disease-causing human DNA variants.16

**Sources of Funding**
This work was supported by National Institutes of Health grant R01 HL108943.

**Disclosures**
None.

**References**
6. Schirle NT, MacRae IJ. The crystal structure of human Argonaute2. *Science*. 2012;336:1037–1040.

Key Words: Editorials ▪ cardiomegaly ▪ microRNAs ▪ RNA
Ménage à Trois: Intimate Relationship Among a MicroRNA, Long Noncoding RNA, and mRNA
Gerald W. Dorn II and Scot J. Matkovich

Circ Res. 2014;114:1362-1365
doi: 10.1161/CIRCRESAHA.114.303786

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circres.ahajournals.org/content/114/9/1362

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation Research can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation Research is online at:
http://circres.ahajournals.org//subscriptions/