From RNAi to Epigenomes: How RNA Rules The World

Eric Westhof*[a] and Witold Filipowicz*[b]

Last October, the Boehringer Ingelheim Fonds sponsored and organized an international conference on RNA Silencing that was chaired by Thomas Tuschl from Rockefeller University, New York, and Thomas Jenuwein from the Institute of Molecular Pathology, Vienna. During two days, twenty-five speakers presented their results to an attentive and very reactive audience. It cannot have escaped the attention of any scientist interested in biology that tiny RNA molecules are transforming our thinking on genetic regulatory processes. It is now realized that post-transcriptional gene silencing (PGTS) or cosuppression in plants, quelling in fungi and RNA interference in animals all share common processing mechanisms.[1–3] RNA interference (RNAi) mechanisms.

RNA interference (RNAi) describes the process by which double-stranded RNA (dsRNA) introduced into cells leads to the degradation of the messenger RNAs, which contain regions homologous to the triggering dsRNA. In 1998, Fire, Mello and co-workers demonstrated that dsRNA promotes the degradation in a sequence-specific manner of endogenous RNAs in a worm.[4] It was already known by then that RNAi/PTGS in plants[5] and worms[6] was systemic and heritable. These two properties require catalytic and amplification mechanisms, which have been partly biochemically characterized.[1–3] Recent experiments also provide some hints as to how RNAi can spread throughout the organism.[7–8]

Starting, in 1999, with the establishment of a cell-free Drosophila system able to recapitulate many of the features of RNAi, Tuschl and collaborators could demonstrate that RNAi can be mediated by sequence-specific processes in soluble reactions.[9] By using this system, it could be further shown that, during RNAi, both strands are processed to 21–23 nt RNAs in an ATP-dependent manner.[10] Strikingly, small RNAs of similar size have been previously identified as molecules that accompany RNA silencing in plants.[11] Simultaneously, a silencing complex containing small RNAs was beginning to be characterized.[12] Processing of dsRNA to small RNAs does not require the target mRNA; however, the cleavage of the mRNA is guided by the pairing to the complementary region. The demonstration that duplexes of 21-nt RNAs with 2-nt 3’ overhangs (small interfering RNAs, siRNAs) are the sequence-specific mediators of RNAi[12] constituted a most crucial discovery in an expanding new field of RNA research. Indeed, soon afterwards, it was demonstrated that siRNAs can be used for the silencing of gene expression in mammalian cells without activation of the mammalian interferon system, thereby providing a new tool for genome-wide analysis of human gene function.[13, 14] For a recent debate about some siRNAs being nevertheless able to induce the interferon response, see the article by Jackson and Linsley.[15] A little bit later, formation of both siRNAs and the genome-encoded ~21-nt RNA regulators, originally known as small temporal RNAs but soon rebaptized microRNAs (miRNAs), was shown to require DICER, a multidomain nuclease related to RNase III enzymes that have specificity for dsRNA.[16, 17]

Genetic, biochemical and bioinformatic approaches identified large numbers of miRNAs and their genes in both invertebrates and vertebrates, thus demonstrating that these tiny RNA regulators are not the oddity of the nematode worm Caenorhabditis elegans, where they were originally discovered in 1993.[18–20] siRNAs in plants and animals associate with a set of specific proteins to form the effector complex, called RISC (for RNA-Induced Silencing Complex), which mediates the mRNA cleavage in the middle of the sequence complementary to siRNA (Scheme 1). In contrast, most miRNAs in animals form imperfect duplexes with sequences in the 3’-untranslated region (3’-UTR) of mRNAs and block protein synthesis by an unknown mechanism (Scheme 2). In plants, however, miRNAs generally base-pair quite precisely with their targets and guide, as do siRNAs, the mRNA cleavage.[18a, b] miRNAs form part of RISC-like ribonucleoprotein particles, miRNPs.[20a, b] The protein composition of RISC and miRNP particles partly overlaps (for example, both types of complexes contain proteins of the Argonaute family); this is consistent with miRNAs’ being able to act as siRNAs under some circumstances, and vice versa.[3, 19, 20]

These RNA-guided silencing mechanisms opened a whole new field not only in RNA research but also for our un-

![Scheme 1. The RNA interference reaction and the complexes mediating mRNA cleavage (RISC) and chromatin silencing (RITS). Small interfering RNAs (siRNAs), generally not conserved in evolution, are processed from double-stranded RNAs derived from aberrant mRNAs, transcripts of transposons and heterochromatic DNA repeats and viruses.](image)
understanding of the basic mechanisms underlying development and function. The potential for therapeutic applications to cure cancer, and genetic and viral diseases is being actively explored. The diverse and multifaceted roles of noncoding RNAs in molecular biology are only being realized nowadays. Clearly, the breadth and complexity of the regulation, from plant to animal cells, were not suspected and still defy our understanding. At the meeting, recent developments in our understanding and biochemical characterization of the catalytic aspects were dissected.

Of most recent findings, the demonstration that one of the four known mammalian Argonaute proteins, Ago2, functions as a RISC catalytic component, cleaving mRNA, is particularly interesting. The groups of Hannon and Joshua-Tor have crystallized the archaeabacterial Ago-like protein and found that one of its conserved domains, Piwi, has a fold similar to that of RNase H, an enzyme known to cut RNA strands in RNA–DNA duplexes. Ago2 mutagenesis, guided by the structural information, showed that its Piwi domain specifically cleaves the siRNA–mRNA duplex. While Ago2 is the only cleaving Argonaute protein, Ago2 mutagenesis, guided by the structural information, showed that its Piwi domain specifically cleaves the siRNA–mRNA duplex.

The seven to eight nucleotides in the 5' segment of the miRNA are most important for the specificity of binding to the target mRNA. Nonetheless, the relaxed pairing requirements for animal miRNAs render in silico predictions of miRNA targets challenging, although they are starting to yield results. However, genetic and biochemical approaches have to date allowed approximately a dozen mRNA targets in different organisms to be identified and validated. The latest additions to the list include miRNAs that regulate neuronal asymmetry in C. elegans, the HOX gene hox88 in mouse embryos, and mammalian myotrophin (Mtpn) mRNA, which functions in insulin secretion. The latter mRNA is regulated by miR-375, specifically expressed in pancreatic endocrine islets. Together with miR-375, ten other new miRNAs specifically expressed in the murine pancreatic β-cell line were identified. Recurrent reports on cloning of novel miRNAs from specific cells or tissues, argue that the number of different miRNAs expressed in mammals might be much higher than originally predicted. Since dozens of genes are thought to be regulated by one miRNA, the fraction of the genome controlled by miRNAs might be very large. Pfeffer et al. found that the Epstein Barr Virus, a member of the Herpes family, encodes its own miRNAs, which probably act as regulators of both host and viral gene expression.

Another topic brought to the forefront in this meeting was the relationships between RNAi and epigenetic mechanisms. Epigenetic mechanisms allow identical genomes to exist in different states, active or silenced, despite cellular division. Chromosomes contain regions that are transcriptionally active, the accessible euchromatin, and other regions which are transcriptionally inert, the condensed heterochromatin. The organization of chromatin is based on the repetitive folding of nucleosome particles in which approximately two turns of DNA wrap around the octamer core histone proteins. The epigenetic marks on the genome are on both the DNA and the histone proteins. Some cytosine bases of DNA are methylated while some amino acid residues (arginines, lysines or serines) of the N-terminal histone tails can be methylated, acetylated or phosphorylated. The hypothesis is that the resulting distribution of modifications, the "histone code," attracts and retains protein effectors; this leads to a variety of "epigenomes".

Now, how is the link between RNAi and heterochromatin silencing achieved? Heterochromatin, located at the centromeres and the telomeric ends of chromosomes, and also in some other regions of the genome, is generally considered as "junk" DNA, containing heavily repeated noncoding DNA sequences and transposable elements, either active or inactive. The key link lies in the nature of the heterochromatin repeats. These are long stretches of DNA sequences frequently arranged as inverted repeats. Transcription through inverted repeats or copying of both strands of DNA leads to the formation of dsRNA or RNA hairpins, the precursors necessary for entering the RNAi pathway (Scheme 1).

Deletions of genes involved in the RNAi machinery, together with the cloning of siRNAs from silenced regions, demonstrated its role in heterochromatin silencing. In addition, a complex comprising siRNAs and the Ago protein, named RITS (RNA-induced Initiation of Transcriptional gene Silencing; Scheme 1), was shown to mediate the process in Schizosaccharomyces pombe. A connection of the RNAi machinery with DNA methylation, chromatin silencing and centromere function had already been found in plants and, more recently, also in metazoa, including vertebrates. Even more intriguingly, a miRNA has recently been identified that guides gene-specific DNA methylation in Arabidopsis by associating with the nascent miRNA transcript in the nucleus. How many more surprises lie ahead of us?

The discovery, in the early eighties, by Sidney Altman and Thomas Cech of catalytic RNAs led to an amazing surge in RNA research, with the development of new techniques and an enormous increase in knowledge about RNA folding.

Scheme 2. The biogenesis and translational inhibitory action of miRNAs. Note that plant, animal and viral miRNAs are also known to direct cleavage of their natural targets. Drosha is an RNase III-like enzyme that initiates processing of miRNA precursors in the nucleus. Dicer then produces, in the cytoplasm, the siRNA-like miRNA duplex, one strand of which (corresponding to mature miRNA) is incorporated into the RISC-like miRNP complex.
and activity. The realization of the catalytic potential of RNA strengthened the hypothesis about the role of RNA at the origin of life four billion years ago. The present results lead us to suspect not only that DNA is a modified RNA, but also that its functions are under the control of tiny RNA molecules. The difficulties and potential of therapeutic applications of RNA interference were also discussed at the meeting. As is often the case, the main obstacle to applications to mammalian cells and organs is delivery of the siRNAs.[17] Another difficulty arises from off-target effects due to cross-binding between introduced siRNAs and messengers with sequence similarities.[18] However, high-throughput RNAi screens have been developed in mammalian cells that have already led to the identification of loss-of-function genes in mammalian cells.[14,59,60] Not surprisingly, the projected worldwide revenues for RNAi as a tool for research in 2010 are estimated at $300 million.[61]

The seeds for this second RNA revolution were planted by scientists fascinated by fundamental problems, the development of a model minuscule worm, flowering in plants, or the catalytic properties of ribozymes. Ironically, some of the crucial developments were achieved because some C. elegans controls did not behave the way they were expected to,[62] or Petunia flowers surprisingly turned white instead of dark purple.[63,64]