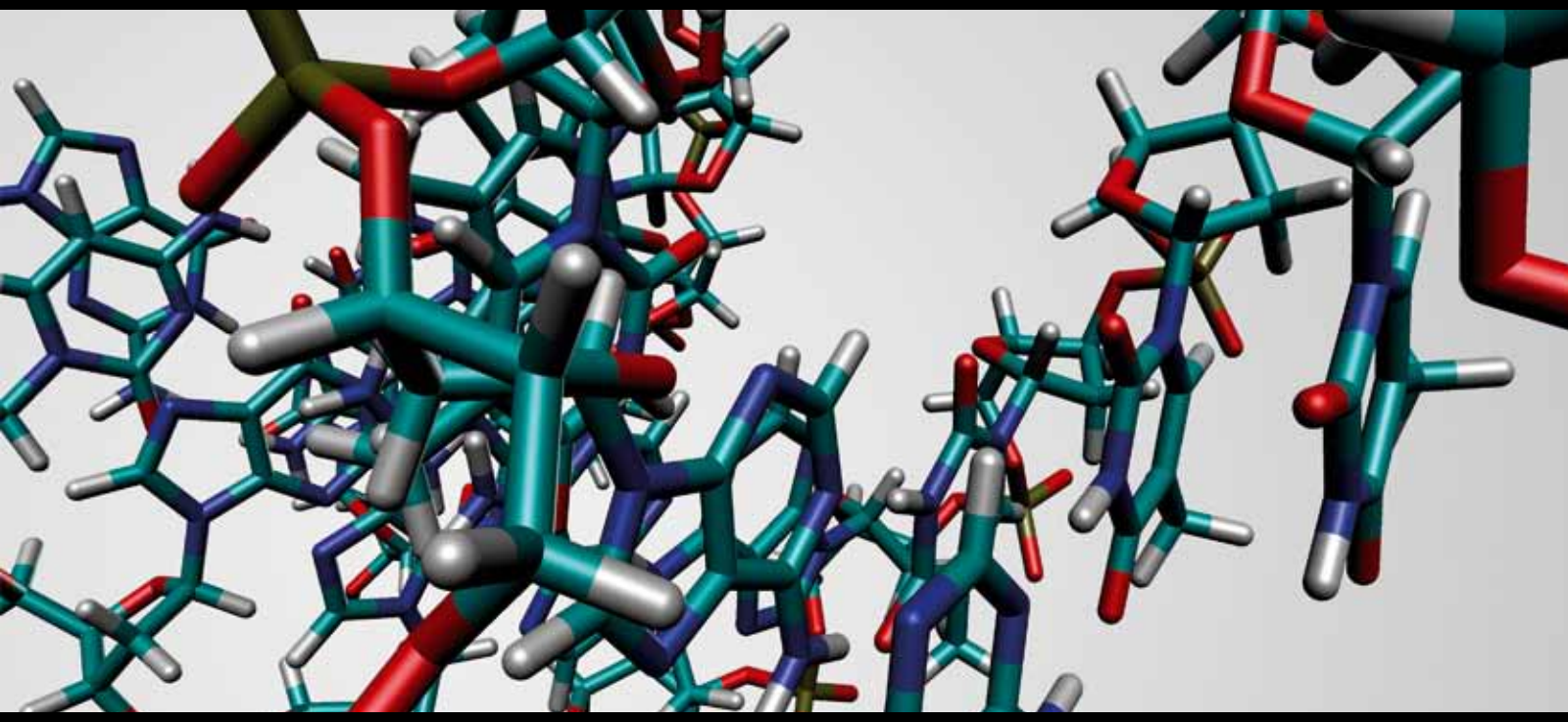
Journal of Nucleic Acids

Using Chemical Approaches to Understand RNA Structure and Function in Biology

Guest Editors: Dmitry A. Stetsenko, and Arthur van Aerschot



Using Chemical Approaches to Understand RNA Structure and Function in Biology

Journal of Nucleic Acids

Using Chemical Approaches to Understand RNA Structure and Function in Biology

Guest Editors: Dmitry A. Stetsenko and Arthur van Aerschot

Editorial Board

Contents

Hindawi Publishing Corporation
Journal of Nucleic Acids
Volume 2012, Article ID 972575, 2 pages
doi:10.1155/2012/972575

Editorial

Using Chemical Approaches to Understand RNA Structure and Function in Biology

Dmitry A. Stetsenko[1] and Arthur van Aerschot[2]

[1] *Department of Bioengineering, Imperial College London, South Kensington Campus, London SW7 2AZ, UK*
[2] *Department of Pharmaceutical Sciences, Katholieke Universiteit Leuven, Minderbroedersstraat 10, 3000 Leuven, Belgium*

Correspondence should be addressed to Dmitry A. Stetsenko, d.stetsenko@imperial.ac.uk

Received 10 August 2011; Accepted 10 August 2011

Interest of the research community in the aspects of chemical biology of RNA has increased vastly over the last twenty years, primarily due to the discovery of RNAi, our deepened understanding of the role of miRNA in the subtle regulation of vital cellular processes, and the realization of the fact that the RNA world—the realm where RNA plays the key parties as a self-replicating molecule and a universal catalyst—is still pretty much with us today as the ribosomal RNA performs its catalytic solo in the formation of the peptide bond in the ribosome. To answer the needs of Biology, Chemistry had, in turn, multiplied and perfected its approaches to study the molecular mechanisms underlying RNA functions in living systems. So, the idea behind this special issue is to show our readership a screenshot of what could be, and has been, achieved recently by applying chemical methods to solve the problems of RNA biology. Ten articles have been carefully selected out of the bunch of those submitted to provide, we believe, a balanced view of different facets of RNA structure and function and, also, of the array of chemical tools to enable us to peek into them.

The issue is a mix of reviews and research papers and thus falls into two parts. We have purposefully incorporated a number of authoritative overviews to paint a wider background of RNA functioning in the first half. It starts with a general treatise on chemical approaches to study RNA structure-function relationships in the postgenomic era by T. S. Ro-Choi and Y. C. Choi. Dependence of RNA-protein interaction on RNA conformation is highlighted by a medicinally important example of a noncoding RNA BC1 and the Fragile X Mental Retardation protein in another paper by X. Yan and R. B. Denman. The one that follows by D. Kurita et al. amply shows how chemical probing has provided insights into the mechanisms of *trans* translation and the role played by transfer-messenger RNA (tmRNA). A Herculean task of the identification of minor nucleotides in RNA and the use of specific chemical reagents for its accomplishment are reviewed in the contribution from a group of French researchers led by Professor Y. Motorin. Taking snapshots of RNA in live cells by means of molecular beacons technology is the subject of a review of R. Monroy-Contreras and L. Vaca.

Then the issue goes into details of the interplay between a single chemical, resveratrol, and miRNAs during inflammatory processes and cancer, reviewed by E. Tili and J.-J. Michaille. The miRNA theme continues in the review by A. Serva et al. of the application of high-throughput functional analysis for elucidation of the multiple biological roles of miRNAs. Here we cross the boundary into the second half of our special issue composed from research papers, and the one that comes next by K. Terazawa et al. opens the RNA interference section with synthetic short hairpin RNAs (shRNAs) as potent RNAi inducers. Another article, commissioned by researchers from the Iberian Peninsula R. Eritja and coworkers, deals with branched RNA architectures for RNAi.

Finally, a paper of the issue done by M. A. Zenkova et al. focuses on an applied aspect of RNA technology, namely, the development of sequence-specific artificial ribonucleases

incorporating multiple imidazole residues that perform similar function to their counterparts in natural RNA-cleaving enzymes.

In conclusion, the issue attempts to present a selection of reviews and research papers, which are designed to illustrate how by using various implements and gadgets from the ever-growing arsenal of Chemistry one can hopefully gain a better understanding of what is happening when RNA molecules interact with their manifold targets and perform their plethora of functions. It must be added that this is by no means a comprehensive picture but rather a momentary screenshot of a dynamic process of harnessing the power of Chemistry to serve the ends of Biology. Sapienti sat.

Dmitry A. Stetsenko
Arthur van Aerschot

Hindawi Publishing Corporation
Journal of Nucleic Acids
Volume 2012, Article ID 369058, 40 pages
doi:10.1155/2012/369058

Review Article

Chemical Approaches for Structure and Function of RNA in Postgenomic Era

Tae Suk Ro-Choi and Yong Chun Choi

Departments of Pharmacology and Biochemistry, College of Medicine, Dong-A University, 3-1 Dong Dae Shin Dong, Seo Gu, Busan 602-714, Republic of Korea

Correspondence should be addressed to Tae Suk Ro-Choi, tsrochoi37@yahoo.com

Received 18 October 2010; Revised 21 June 2011; Accepted 5 July 2011

Academic Editor: Dmitry A. Stetsenko

In the study of cellular RNA chemistry, a major thrust of research focused upon sequence determinations for decades. Structures of snRNAs (4.5S RNA I (Alu), U1, U2, U3, U4, U5, and U6) were determined at Baylor College of Medicine, Houston, Tex, in an earlier time of pregenomic era. They show novel modifications including base methylation, sugar methylation, 5′-cap structures (types 0–III) and sequence heterogeneity. This work offered an exciting problem of posttranscriptional modification and underwent numerous significant advances through technological revolutions during pregenomic, genomic, and postgenomic eras. Presently, snRNA research is making progresses involved in enzymology of snRNA modifications, molecular evolution, mechanism of spliceosome assembly, chemical mechanism of intron removal, high-order structure of snRNA in spliceosome, and pathology of splicing. These works are destined to reach final pathway of work "Function and Structure of Spliceosome" in addition to exciting new exploitation of other noncoding RNAs in all aspects of regulatory functions.

1. Introduction

A key element in the study of cellular RNA metabolism is the molecular characterization of RNA. This characterization requires accurate determination of the RNA sequence. It is imperative to understand how RNA structure complements the functional definition of RNA. Cellular RNAs are posttranscriptionally modified at various points in the primary RNA transcript as well as processed. In cellular RNA metabolisms, RNA maturation is performed through various structural alterations that include chemical modifications of constituent components. A most representative modification is observed in chain shortening, rearrangements by transfer of phosphodiester linkages involved in splicing mechanisms (pre-mRNA), deletions (pre-rRNA), and transsplicing (trypanosomal mRNA). Another is chain expansion demonstrated by modifications observed on polyadenylation, U-addition at 3′ ends, 5′-cap formation at 5′ ends, and insertions within trypanosome RNA. Other examples of modifications are base modifications, such as deaminations, methylations, hypermodifications, and ribose methylations.

The most modified RNAs are tRNAs containing approximately 2–22 modified nucleotides per molecule of ~75 nucleotide length, and there have been more than 130 different signature modified nucleotides reported [1]. The discovery of snRNA and $m_3{}^{2.2.7}G$ caps occurred within the last 50 years. They also contain their own specific modified nucleotides such as Ψ, m^6A, m^2G, and 2′-O-methylated nucleotides (Table 1).

The next class is the ribosomal RNAs which contain 204–209 modified nucleotides within 18S (1,869 nt) + 28S (5,035 nt) RNA in eukaryotes. The mRNAs contain the least modified nucleotides, with the exception of the 5′ end cap structure and occasional m^6A in the molecule.

In ensuing years, massive scale DNA sequencing was advanced to accommodate the "Human Genome Project." Two groups published the genomic map where the coding genes were cataloged. It was conservatively estimated that there are 25,000 genes and 50,000 proteomes involved in cell metabolism. It was also envisioned that processing mechanisms could be discerned by comparing the genomic structure with the RNA sequence determined using cDNA methods. Based

Table 1: Signature sequences and modifications of major snRNAs. The 5′ cap and 3′ nucleosides, base modified nucleosides, and alkali resistant oligonucleotides were determined by many methods described in the text. The table provides a summary of individual RNA characteristics of rat Novikoff hepatoma cells.

RNA	Subspecies	Localization	5′ End	3′ End	Modified Nucleotides
4.5S RNA	I	Extranucleolar nuclei	pppG	U-OH	
	II (U6)	Extranucleolar nuclei	mpppG	U-OH	3Ψ, m^6A, m^2G, AmA, AmG, AmGmC, CmC, CmA, CmCmU
	III	Extranucleolar nuclei	pA	Um	AmA, GmA, 2GmG, m^2G, m^6A, 3Ψ
5S RNA	I	Nucleoli, nuclei and cytoplasm	pppG	U-OH	
	II	Nucleoli, nuclei and cytoplasm	pppG	U-OH	
	III (U5)	Extranucleolar nuclei	$m_3^{2,2,7}$GpppAmUmAC	U-OH A-OH	UmU, GmC, 2Ψ
U1 RNA	U1a (5.8S RNA)	Nucleoli and cytoplasm	pC, pG	U-OH	UmG, GmC, Ψ
	U1b, U1c	Extranucleolar nuclei	$m_3^{2,2,7}$GpppAmUmAC	U-OH, G-OH	AmC, 2Ψ
U2 RNA		Extranucleolar nuclei	$m_3^{2,2,7}$GpppAmUmC	C-OH A-OH	GmGmC, GmG, GmA, m^6AmG, CmΨ, UmA, CmU, 13Ψ, (m^6A, m^2G)
U3 RNA	U3a, U3b, U3c, U3d	Nucleoli	$m_3^{2,2,7}$GpppAmA(m)AG	A-OH, U-OH, C-OH	2Ψ

on the ever-increasing number of RNA sequences, it was determined that most coding RNAs mature as a result of alternative splicing. Aberrant splicing is attributed to point mutations in the genetic code and splicing code [2]. It is noted that RNA sequencing can aid the determination of the molecular pathogenesis of diseases.

2. Historical Venture of RNA Research

Detailed nucleic acid chemistry began with discoveries of the DNA helix by Watson and Crick [3] and DNA polymerase by Lehman et al. [4, 5]. With DNA being the genetic material providing a blue print for living creatures, it moved genomic era thinking away from the earlier notion that protein, carbohydrate, and lipid were the only essences of living things.

DNA is there to provide information needed to build the cells, tissues, organs, and whole individuals. It took a long time to move from the histochemical presentation of DNA in the nucleus and RNA in the nucleolus and cytoplasm [14] to the isolation of nucleoli, nuclei, mitochondria and ribosomes, facilitating the elucidation of their components, their structures, and their functions. Even within the same species, no two individuals are identical. Disarray in DNA structure can determine whether one is healthy or diseased. In the quest to conquer cancer, differences in cellular morphology and uncontrolled growth became and remain a major research consideration when one compares normal cells with cancerous cells and tissues. Cancer cells with pleomorphic, hypertrophic nuclear, and nucleolar morphology remain a useful pathological criterion for a cancer diagnosis. The information within genes is transferred to RNA and then to proteins made on ribosomes that define a cell phenotype. The fractionation of cells into various components includes nucleoli, nuclei (Figure 1), ribosomes, mitochondria, cytosol and others.

The main interest among these compartmental components was the RNA. The RNA has its own exclusive properties which are not found in DNA.

The discovery of RNA polymerase I in the nucleoli [31] is the landmark of RNA research in these cellular compartments. It was not until 1968, with the introduction of gel electrophoresis into RNA research [32], that subspecies of 4–8S RNAs could be separated from high-molecular-weight RNAs (>18S RNA). Until then, the 4–8S RNAs were considered as tRNAs and their precursors. Different from the prokaryotic cells, eukaryotic cells were shown to have a variety of small RNAs in their nuclei (Figure 2). These RNAs used to be called LMWN RNA (low-molecular weight nuclear RNA) and now the name is unified as snRNA (small nuclear RNA).

These include U1 RNA, U2 RNA, U3 RNA, (named as such because these RNAs contain a high proportion of uridylic acid), 5S RNA III (U5 RNA), 4.5S RNA I (Alu RNA), 4.5S RNA II (U6), and 4.5S RNA III. All of these snRNA species and many more have been sequenced and their functions elucidated in pre-rRNA processing [33] and pre-mRNA splicing [34, 35].

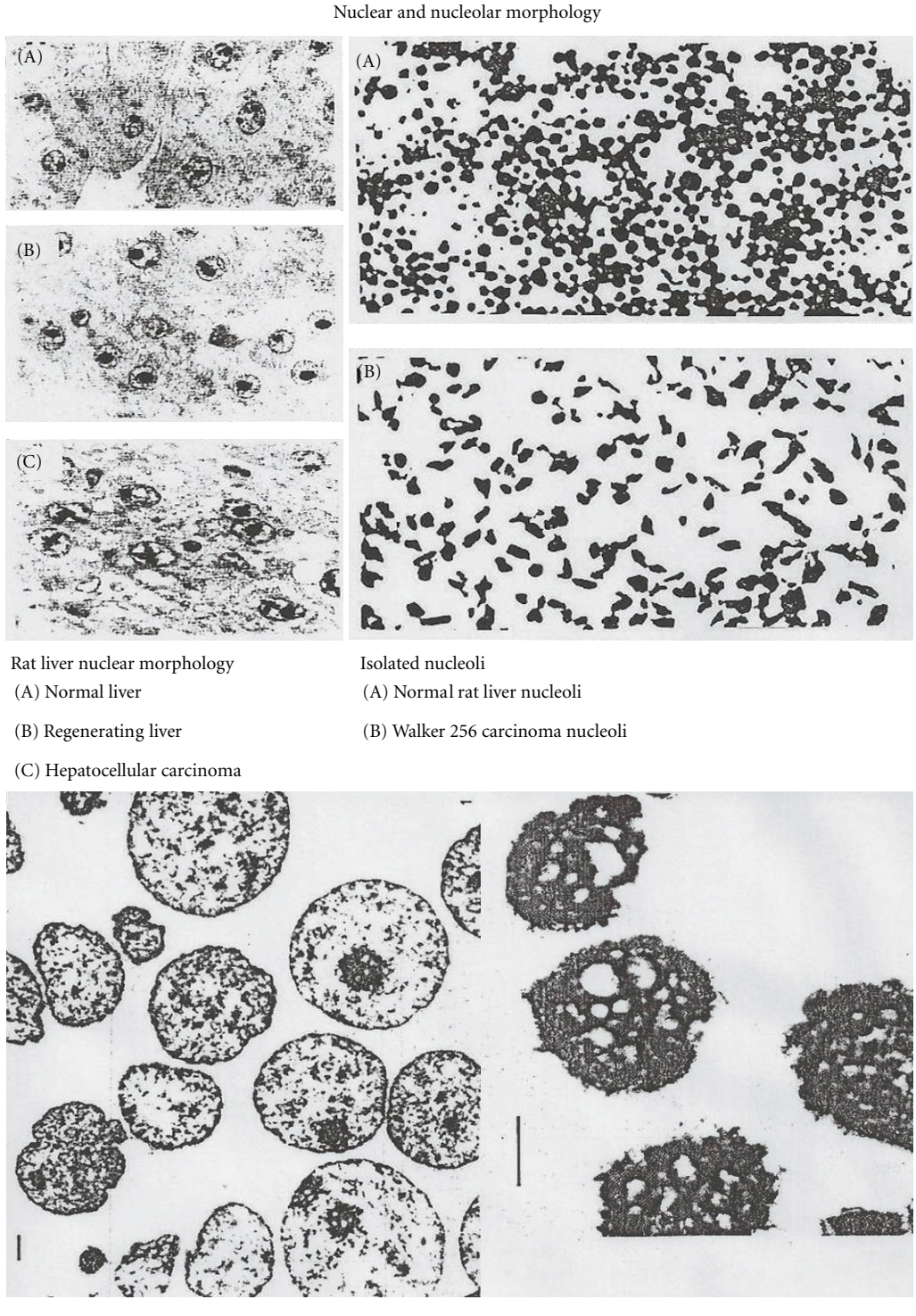
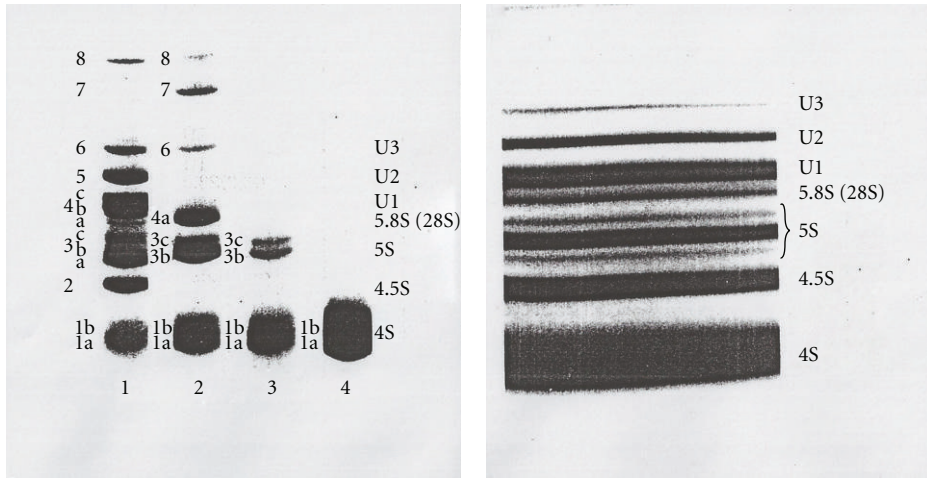


FIGURE 1: Nuclear and nucleolar morphology. Normal rat liver nuclei have 1–6 round nucleoli which are less than 2 μm in diameter. In regenerating liver, cells contain enlarged nucleoli. In tumor cells (hepatocellular carcinoma), the nucleoli are not only enlarged but also they become pleomorphic in morphology. Nuclei were isolated by homogenization in 2.3–2.4 M sucrose containing 3.3 mM $CaCl_2$. Nuclei were sonicated in 0.34 M sucrose and layered on 0.88 M sucrose for purification by centrifugation. Isolated nuclei and nucleoli had high purity, and morphologies were well preserved [6].



(a) (b)

Figure 2: Polyacrylamide gel electrophoretic separations of nuclear 4–7S RNAs of rat liver and Novikoff hepatoma cells [7]. (a) The 8% gel electrophoretic patterns of 4–7S RNA from various cell organelles of rat liver. The gel was stained with methylene blue for RNA visualization. (1) Nuclear 4–7S RNA, (2) ribosomal 4–7S RNA, (3) mitochondrial 4–7S RNA, and (4) soluble cytoplasmic sap 4–7S RNA. (b) The 10% slab polyacrylamide gel electrophoretic separation of [^{32}P]-labeled 4–7S RNA from Novikoff hepatoma cell nuclei. The gel was autographed with X-ray film.

Table 2: Characteristics of RRM/RNP/RBD domain.

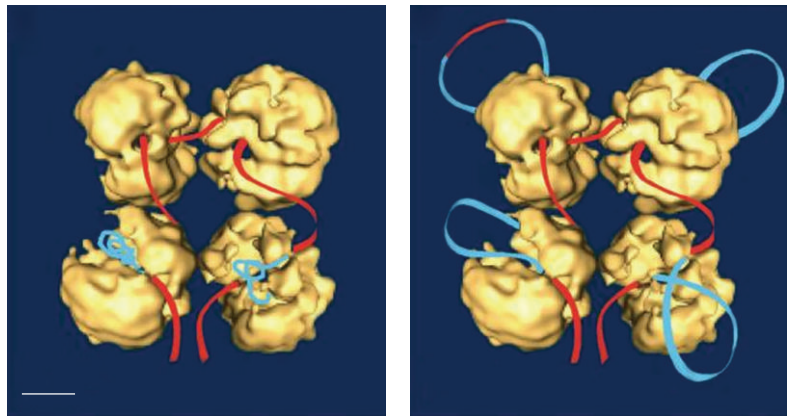
(1)	~90–100 amino acids domain and most abundant in vertebrates
(2)	Many RNA binding proteins contain more than one RRM
(3)	Contain 2 conserved RNP1 (RGQAFVIF in β3) and RNP2 (TIYINNL in β1) in 4 antiparallel β-sheets of $\beta\alpha\beta\beta\alpha\beta$-fold
(4)	Binds 2–8 nucleotides of RNA (2 in CBP20, nucleolin and 8 in U2B″)
(5)	A typical RRM containing 4 nucleotide binding sites (UCAC)
(6)	3 conserved aromatic amino acids (Y, F, W, H or P) in central β-strands (2 in RNP1 of β3 and 1 in RNP 2 in β1)
(7)	2 RRMs in a protein are separated by small linker and provide a large RNA binding surface or RNA binding surface point away from each other
(8)	RNA bases are usually spread on the surface of protein domains while the RNA phosphates point away toward the solvent
(9)	Binding surface of the protein is primarily hydrophobic in order to maximize intermolecular contact with the bases of the RNA
(10)	Few intramolecular RNA stacking and many intermolecular stacking mediated by aromatic amino acids
(11)	RNA recognition is a two-step process, in which any RNA is attracted approximately equally well. However, if stacking and hydrogen-bond interactions that "lock" the interaction cannot be properly established, the complex redissociates quickly (large k_{off}), which results in overall weak affinity for RNA oligonucleotides of the wrong sequence
(12)	Many ssRNA binding proteins recognize RNA in the loop (stem-loop) better than in ssRNA (k_{on} ~ 3 fold & k_{off} ~ 590 fold, therefore, overall affinity ~2000 fold differences) due to higher entropy loss with ssRNA binding than stem-loop binding and stabilizing interactions of stem

The most interesting discoveries in the midst of sequencing were the very unusual trimethylguanosine cap structure in U1 RNA ($m_3^{2,2,7}$GpppAmUmAC), U2 RNA ($m_3^{2,2,7}$GpppAmUmC), U3 RNA ($m_3^{2,2,7}$GpppAmA(m)-AGC), and 5S RNA III (U5 RNA) ($m_3^{2,2,7}$GpppAmUmAC) [36]. Afterwards, myriads of cap structures in viral RNA and mRNA were discovered [37].

The history of RNA sequence work has occurred in three eras. The pregenomic era was devoted to the small RNAs and commenced with the sequence of large RNAs as technology developed for cDNA synthesis, amplification, cloning, and sequencing. The DNA technology was explosive and paved the way toward establishment of sequence technology not only for RNA and cDNA but also for genomic DNA.

In addition to sequence study, the secondary and tertiary structures have also been determined. A representative study was the crystallographic study of RNA-protein interactions. For example, the most well-worked-out motif is RRM (RNA recognition motif) which is most abundant in hnRNP [40] and splicing factors [41]. The summary of characteristics of RRM is in Table 2.

It has been known for a long time that pre-mRNA (hnRNA) is cotranscriptionally assembled into beads on a string consisting of 30–50S (20–30 nm) particles [42]. The



(a) (b)

Figure 3: The supraspliceosome model from the article by Sperling et al. [8]. (a) It was stated that pre-mRNA which is not being processed is folded and protected within the native spliceosome. (b) With different staining protocol, it was possible to visualize the RNA strands and loops emanating from the supraspliceosome. These complexes were found to contain hnRNP proteins (personal communication).

RNP (hnRNP) has usually 48 hnRNP proteins and ~700–800 nucleotide long RNA string [43]. More recently, most hnRNP proteins have been found to have 1-2 RRM motifs for RNA binding. From these characteristics, the primary RNA transcripts have been folded from the 5′ end with the following rules: a minimum of 3 nucleotides in the loop and a minimum of 3 base pairs at the stem. According to stacking and loop energy rules, two nucleotide loops cannot exist. The number of base pairs needed for stabilization with the most stable stacking energies by CCC/GGG or GGG/CCC is 3 base pairs with −9.8 kcal and the highest loop destabilizing energy is +8.4 kcal [44]. In addition, protein binding to RNA has been shown to have $-\triangle G \approx 10-13$ Kcal/mol [45] which can overcome the loop destabilizing energies of any size. With this rule, folding the hnRNA in GC, AU, and GU pairings was carried out as the RNA was transcribed, extending contiguous base pairing until it comes to a base pair mismatches. Accordingly, small simple RNA hairpins have been constructed with the aid of a computer [46] from the 5′ end (transcription start sites). Consensus patterns for folding characteristics have been observed (Table 3).

The transcripts form one stem loop for every 15–18 nucleotides which is consistent with ~15–17 nucleotides per hnRNP protein (700–800 nucleotides per 48 hnRNPs in one hnRNP particle) reported earlier [43]. The thermodynamics of RNA folding was consistent with the order of splicing in ovomucoid pre-mRNA [47]. From the point of view that supraspliceosomes contain hnRNP proteins (personal communication), it may be that this cotranscriptional formation of hnRNP string particles [47–49] may contribute to a role in the formation of supraspliceosomal RNP (Figure 3) [8].

The postgenomic era is the present day era or the second generation genome era. With the recent discovery that there is a paradox [50, 51] in the cellular transcript number, which is 2-3-fold in excess and that 50% of the cellular transcripts are ncRNAs, the second generation genomic era is in the process of resequencing the genome for ncRNAs. It is anticipated that there will be a revision in the first generation genomic picture. In this era, work is proceeding that will probe and dissect the RNA metabolism in which aberrant processing should be elucidated by RNA sequencing. To dissect the molecular pathology of RNA metabolism, it is also necessary to study higher-order structures based on the sequence studies involved in the assembly of macromolecular machinery. It is natural to hope that therapeutic interventions will be discovered that can correct errors in the genetic code and its product splicing.

Table 3: Frequency of stem loops in primary pre-mRNA transcripts. The simple stem loops with minimal 3 nucleotides in the loop and minimal 3 base pairs in the stem consisting of AU, GC, and GU pairs have been constructed with the aid of a computer [46]. The total number of nucleotides were divided by numbers of stem loops for frequency. The number of nucleotides in each loop and each stem and spacer were counted and averages were calculated. (1) Human insulin gene transcript: 1,430 nt. (2) Human HDHGT (25-hydroxyvitamin D3 1-α-hydroxylase gene transcript): 4,825 nt. (3) Human FMR1 (fragile mental retardation 1) gene transcript: 39,224 nt. (4) Chicken ovomucoid gene transcript: 6,067 nt.

Transcript	nt/loop	nt/stem	nt in spacer	Frequency
(1) Insulin	4.6	7.4	3.5	15.5
(2) HDHGT	5.8	7.0	4.9	17.6
(3) FMR1	5.0	6.8	3.4	15.3
(4) Ovomucoid	5.6	7.0	3.7	16.0

Table 4: Paradoxical characteristics of ncRNAs in humans and mice [50, 51]. The excessive number of transcripts than anticipated for 25,000 genes indicates that the ncRNAs which were not detected due to scarce abundance have been detected by more sensitive methods. Some of these characteristics are summarized.

	Human	Mouse
Gene Number	75,000, 84,000 or 140,000 (cDNA identified)	
Transcripts		181,000
Population	50% Poly-A RNAs (of 16% genome)	50% transcripts (of 62% genome) (35% from antisense strand)
Intron	30% genome	
Processing	Polyadenylation, 5′ cap, splicing, nucleotide modification	
Transcripts from	Intergenic, Intronic regions and antisense strand	
Short ncRNAs	miRNA, siRNA (tasiRNA, natsiRNA), piRNA, rasiRNA (pitRNA), PARs (PROMTs, PASRs, TSSa-RNAs, tiRNAs), MSY-RNA, snoRNA, sdRNA, moRNA, tel-sRNA, crasiRNA, hsRNA, scaRNAs, AluRNA, YRNA, tRNA-derived RNAs	
Long ncRNA (lncRNA) (0.5–100 kb)	Cancers, disorders in skin, heart, brain, cerebellum, and so forth. TR/TERC, NEAT RNA (NEAT1v-1, NEAT1v-2, NEAT2/MALAT1), PINC RNA, DD3/PCA3, PCGEM1, SPRY4-1T1, xiRNAs (Xist RNA, Tsix RNA, RepA RNA), AIR, H19, KCNQ1ot1, HOTAIR, BORG, CTN RNA, ANRIL RNA, LINE, CSR RNA, satellite DNA transcripts and so forth	
Function	Regulatory function in all aspects of metabolism [52]	

The RNAs have been classified according to the following diverse basis of criteria:

(i) cell biology: cell types, subcellular origins,

(ii) molecular weight: high molecular weight (HMW) and low molecular weight (LMW/small),

(iii) S value: 5S rRNA, 7S RNA, 18S RNA, and others,

(iv) linearity: linear, cyclized, and branched (Y shaped),

(v) metabolism: precursor, processed intermediates, and mature,

(vi) standard: hnRNA, rRNA, mRNA, tRNA, and ncRNA (snRNA, snoRNA, miRNA, and others as in Table 4).

3. Preparation of RNA from Isolated Subcellular Compartments

RNA can be extracted from purified nucleoli, nuclei, ribosomes, mitochondria, and cytosol by the SDS-phenol procedure. The procedure involves the suspension of organelles in 0.3–0.5% SDS (sodium dodecyl sulfate), 0.14 M NaCl, and 0.05 M sodium acetate buffer at pH 5.0 and deproteinization by phenol containing 0.1% 8-hydroxyquinoline at 65°C [53]. The extracted RNA is precipitated with 2–2.5 volumes of ethanol containing 2% potassium acetate. The RNA is washed by ethanol and dissolved in appropriate buffer for the analysis. The DNA and protein contaminations are less than 3% by weight. The purified RNA is separated into individual RNA species using sucrose density gradient centrifugation, gel electrophoresis, and column chromatography [38].

4. Structure Determination

4.1. Structural Characteristics of Various RNAs Bearing Signature Sequences and Modifications. The RNA is composed of basic 4 nucleosides of guanosine, adenosine, uridine, and cytidine linked by 5′-3′ phosphodiester bonds between two ribose moieties. In addition, some of these nucleotides are modified in base as well as in ribose moieties and contain unusual pyrophosphate bonds at their 5′ ends and 2′ O-methylated 3′ end.

Mature RNAs are synthesized in the nuclei and directed by the posttranscriptional processing machineries. Because of these specific modifications, there is a general consensus on the presence of specific signature sequences and modifications for the identity of RNA classes. Based on extensive sequence work, it is possible to classify RNAs according to structural modifications. Figure 4 provides an outline for characteristics of RNA, and its modifications and brief examples are given in Table 5.

4.2. General Scheme of RNA Sequencing. The very first RNA sequence was obtained from the work of yeast alanine tRNA in 1965 [54]. In this work, the prerequisites for RNA sequence work were developed and described. Since then, it is a fundamental approach to establish oligonucleotide catalogs using specific RNases. One set is the catalog of T1 oligonucleotides produced by RNase T1. The other is the catalog of oligonucleotides produced by RNase A. The analytical method was based on UV spectral absorption in the earlier years. Subsequently, since 1970, isotopic labeling methods were widely used which are 1,000-fold more sensitive. Furthermore, many other improvements in RNA

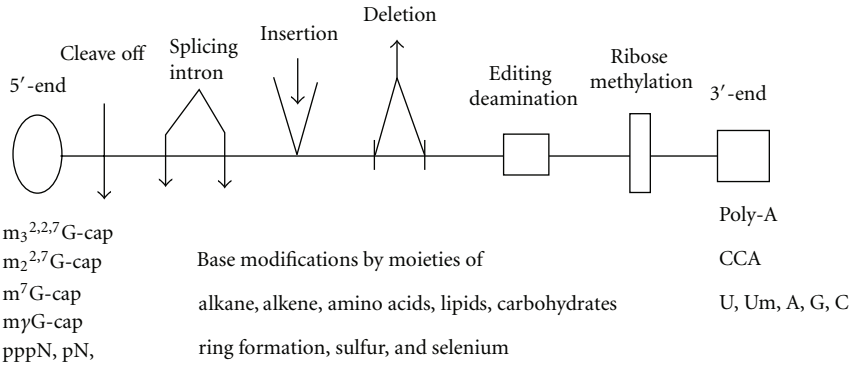


Figure 4: Summary of RNA modifications. Cotranscriptional and posttranscriptional RNA modifications are summarized.

Table 5: Examples of modification. All RNA species including high- and low-molecular-weight RNAs have their own signature sequences and modifications.

RNA	Sequence and signature modification
hnRNAs (Exons + Introns)	m^7G cap, m^6A, Poly-A, splicing codes
mRNAs (mainly exons)	m^7G cap, m^6A, Poly-A Editing (C → U & A → I)
tRNAs	TΨC, CCA, Many hypermodified bases
45S pre-rRNA	Repeated U sequences at 5′ end spacers
18S rRNA	$m_2{}^6$A, hypermodified $m^1acp^3\Psi$
28S rRNA	NmNmNmN, NmNmN
snRNAs, snoRNAs	$m_3{}^{2,2,7}$G cap, mγG cap Types I, II, and III caps by ribose methylations
mRNA (Trypanosome)	$m^7Gpppm_2{}^{6,6}$AmpAmpCmpm^{3}UmpAp Insertion of repeated U sequence Deletion of U sequence

sequence technique have made it possible to advance the rate of RNA sequence work greatly (Table 6).

Improvement was observed in the following areas: (1) RNA labeling techniques, (2) fractionation procedures (chromatography, electrophoresis, and gel procedures), (3) use of various RNases, (4) contig seeking, and (5) ladder sequence gel analysis. For example, based on labeling at the 5′-end with [^{32}P]-γ-ATP by polynucleotide kinase [56], it has become feasible to read a 150 nucleotide sequence using an endonuclease assisted ladder gel from the 3′-end. Also, based on labeling at the 3′-end with [^{32}P]-5′-pCp by RNA ligase [57], it has become feasible to read approximately 150 nucleotides from the 5′-end. Together, these enhancements make it readily feasible to sequence RNA with approximately 300 nucleotides. In contrast to success in the sequence work for small RNAs, two challenges remained. One challenge is related to RNA size and the other is concerned with scarce abundance of RNA in the cell. With the discovery of reverse transcriptase, heat stable DNA polymerase, and recombinant technology, it became possible to produce cDNA, amplify, and clone by RT-PCR methods.

With high-efficiency RT-PCR, high-molecular-weight RNA with 10,000 nucleotides in length can be readily sequenced [59]. A remaining shortcoming of this approach is the inability to fully characterize modified nucleotides. However, ability to deal with long chain lengths and scarce abundance outweighs this limitation. cDNA-based methods clearly dominate any RNA sequence work that involves long RNA length or low RNA abundance. Examples are observed in the direct gene isolation for cleavage controlled processing RNAs (Pre-rRNA and rRNA) and cDNA method for pre-mRNA and mRNAs. Therefore, as a result of accumulated methodologies, it becomes common that RNA sequence can be obtained through more than one scheme or type of technique, such as straight chemical approaches [60] or biotechnology-mediated approaches.

4.3. Outlined Steps of Sequence Work. Brief outlines are described for sequencing RNAs. It may be divided into two methods although combined methodology is in fact feasible.

4.3.1. Direct Method of RNA Sequencing

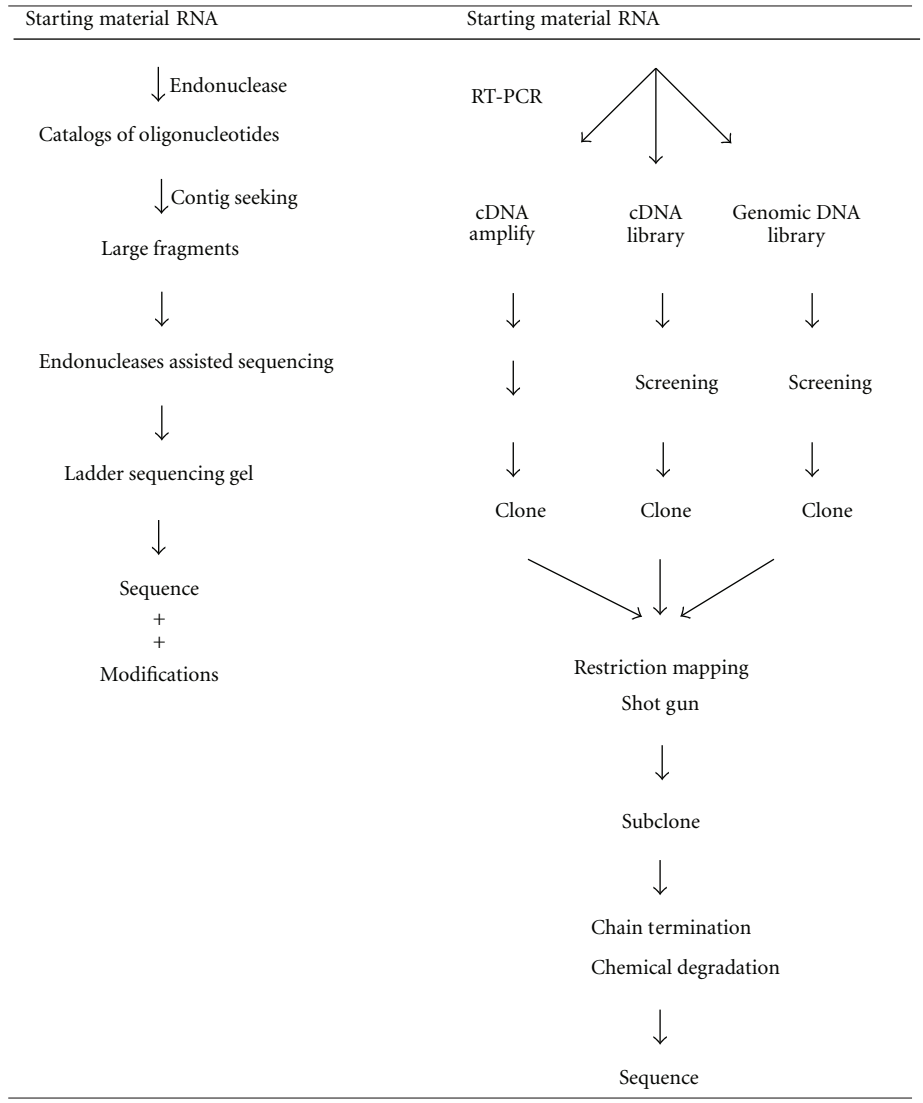
(a) Preliminary Examination of External Glycol Structures. In some cases, a rapid diagnostic examination is required. Most convenient procedures employ the use of specific antibodies against different forms of 5′-cap structure (m^7G cap or $m_3{}^{2,2,7}$G cap) and a oligo-dT column for poly-A affinity chromatography. Alternatively, a [^{3}H]-derivative method can be useful. The radioactive labeling of terminals was performed using the periodate oxidation method, followed by reduction with [^{3}H]-borohydride. T_2 RNase digestion and fractionation by paper chromatography reveal the presence of the 3′-terminal and 5′-cap.

(b) Selection of Labeling Methods. RNA can be labeled in vivo (prelabeling) or in vitro (postlabeling).

In vivo labeling is carried out by incubation of living cells in the presence of [^{32}P]-phosphate in a phosphate-free medium. RNA is uniformly labeled by this method.

In vitro labeling is called postlabeling because it labels the isolated RNA with isotopic agents such as [^{32}P]-phosphate or [^{3}H]-borohydride. [^{32}P]-labeling can be carried out using kinase enzymes. The 5′-labeling is done with [^{32}P]-ATP by

Table 6: General schemes of RNA sequencing. The direct and indirect methods of RNA sequencings are briefly outlined. The cDNA and DNA pathways are considered indirect methods.



polynucleotide kinase, that is, provided the 5′-end is free from phosphate. If the 5′-end is blocked by the presence of a 5′-cap structure, the pyrophosphate moiety must be removed by a pyrophosphatase and phosphatase. And then the kinase method can be employed to introduce the tracer. Labeling at the 3′-end is done with [^{32}P]-pCp by RNA ligase. The [^{3}H]-derivative (nucleotide diol) with [^{3}H]-borohydride indicates that the 3′-end is free from phosphate or any other blocking structures. A shortcoming of [^{32}P]-labeling is the short half-life of the isotope which provides a working period of approximately 4 half-lives. The main limitation of the [^{3}H]-labeling method is weak energy of the tritium isotope. This can make the reading of the autoradiograph for a ladder sequencing gel very difficult.

(c) Initial Reading of Sequence by Ladder Sequencing Gel. To obtain the nucleotide sequence of RNA quickly without characterization of modified nucleotides, it is common to use the endonucleases-dependent sequencing technique [61]. Terminal labeled RNA (5′-end or 3′-end) is partially digested with specific endonucleases (T_1, U_2, A, phys I, and others), and each product is loaded in parallel on a 10–15% denaturing polyacrylamide gel. Note that if crude acrylamide is used, the running temperature of the gel can quickly rise to 60–70°C. Since the mode of cleavage is known, it is possible to discern G (T_1), A (U_2), U and C (A) and C-resistance (Phys I). It is not uncommon to read an RNA sequence using this method within one day.

(d) Base Composition. There are two technical approaches that can be used to determine RNA base composition (levels of nucleotides or of nucleosides).

RNase T_2 or alkali (0.3 N KOH) is used to complete hydrolysis. But alkali (0.3 N K/NaOH) is not preferred because

it destroys 7-methyl purines. Prelabeled [^{32}P]-RNA is hydrolyzed, and its products are separated by 2-dimensional paper chromatography followed by autoradiography [62]. Since the standard separation pattern is known, various modified nucleotides are readily identified by comparison [56].

Alternatively, after cold RNA is digested into constitu-ent nucleotides, which are subsequently dephosphorylated by phosphatase, the resulting nucleosides are converted into [^{3}H]-derivatives and separated by thin layer chromatography. The separated nucleosides (including all modified nucleosides except 2′-O-methylated nucleosides) are detected by fluorography and identified based upon a standard migration pattern (Figure 5) [9].

(e) Catalogs of Oligonucleotides. Two types of catalogs are made. One is an RNase T_1 catalog, and the other is an RNase A catalog.

To map oligonucleotides, two necessary procedures are essential. The first is to prepare labeled oligonucleotides and the second is to fractionate two-dimensionally.

To obtain labeled oligonucleotides, three approaches are possible.

(1) Use of prelabeled [^{32}P]-RNA for specific endonuclease digestion.

(2) 5′ labeling after enzyme digestion using [^{32}P]ATP and polynucleotide kinase.

(3) 3′ labeling after endonuclease digestion and removal of resultant 3′-phosphate by phosphatase. Then the labeled derivatives can be formed by [^{32}P]-5′-pCp and RNA ligase or periodate oxidation followed by [^{3}H]-borohydride reduction.

To Map Oligonucleotides. There are a number of different techniques. However, the most common are a combination of high voltage paper electrophoresis on cellulose acetate at pH 3.5 and high voltage DEAE paper electrophoresis (7% formic acid) or high voltage electrophoresis on cellulose acetate at pH 3.5 followed by DEAE homochromatography at 60–70°C. Another method that can be used is two-dimensional thin layer (PEI) chromatography using two-solvent systems [63]. Detection is performed by autoradiography. It is notable that T1 oligonucleotides from 45S pre-rRNA can be fractionated into approximately 200 spots by homochromatography [64].

To Sequence Oligonucleotides. Several enzymatic digestions can be exploited.

The recovered [^{32}P]-oligonucleotides (prelabeled) are subjected to secondary digestions with RNase U_2 for placement of A residues, RNase T_1 for G residues, RNase A for U, and C residues plus other endonucleases. Treatment with exonucleases (spleen phosphodiesterase, snake venom phosphodiesterase), and partial digestion with the enzymes above is required to sequence RNA. In each step, nucleotide composition is determined.

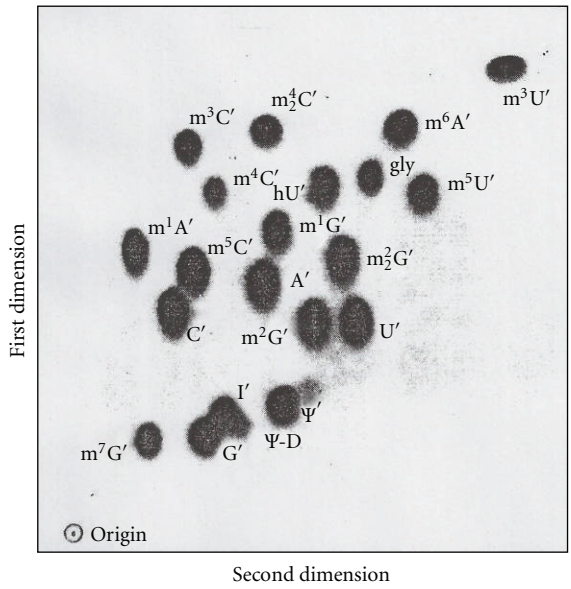


Figure 5: Two-dimensional map of standard nucleosides [9]. The [^{3}H]-labeled standard nucleoside derivatives are separated by two-dimensional thin layer chromatography. The first dimension is shown from bottom to top and the second dimension from left to right. Solvent systems are in the text. N′ represents trialcohol derivatives of representative nucleosides.

To Determine the Sequence of 5′-Labeled [^{32}P]-Oligo-Nucleotides. A mobility shift test can be applied [56]. After partial hydrolysis with snake venom phosphodiesterase the product is fractionated by homochromatography or PEI thin layer chromatography. The mobility shift pattern is produced according to the step-wise loss of each nucleotide from the 3′-end. The resulting pattern can be used to read the sequence of the oligonucleotides.

To Determine the Sequence of [^{3}H]-Oligonucleotides . The procedures used for prelabeled [^{32}P]-oligonucleotides are applicable. Secondary digestion methods and accompanying [^{3}H]-derivative methods for the determination of nucleotide composition can be carried out.

It may be necessary to strengthen the catalog of oligonucleotides. Generally this involves the expansion of the catalog to provide contiguous overlapping sequences. A feasible approach is to produce large fragments (purified on 10–15% denaturing polyacrylamide gel electrophoresis) and identify the overlapping oligonucleotides. Usually a limited fragmentation by a diluted endonuclease at low temperature or water hydrolysis may produce large overlapping fragments [63]. Examination of large fragments, as done above for ladder gel sequencing and catalogs, can often clarify any ambiguity encountered. An excellent example of one hit hydrolysis is observed in the work on tRNA structure [63]. Based on these very same methods, it can be summarized that many small RNAs have been sequenced. These include tRNAs, pre-tRNAs, 4.5S RNA I, 5S rRNA, 5.8S rRNA, snRNAs, snoRNAs, 7S RNA, and some fragments of pre-rRNA, 28S rRNA, and 18S rRNA.

4.3.2. Indirect Method of RNA Sequencing. The indirect method of RNA sequencing using cDNA or DNA gene analysis was developed as part of explosive advancements with DNA biotechnology. The direct RNA sequencing method proved useful for the characterization of small RNAs (~100–300 nt). However, sequencing high-molecular-weight (HMW) RNAs proved to be too difficult. Moreover, HMW RNAs that are scarce abundance often do not meet the sample amounts required by the former methods. The search for a solution to this dilemma was successful. One solution involved the isolation of the gene that codes for a specific RNA and the other is to synthesize cDNA which can also be used to isolate a specific RNA gene. Using DNA biotechnology, it proved possible to scale up and solve "The Human Genome Project." Several genomes have been sequenced, specifically the human (2.9 Gb) and mouse (2.5 Gb) genomes [65–67]. In well equipped laboratories, it is possible to sequence DNA at the rate of 10^6–10^7 nt/day. This technology has been widely commercialized and is currently available as kits for cDNA cloning, sequencing, along with enzymes and equipment that supports automatic sequencing. The principal objective of the genomic approach was to determine the sequences of the coding genes. Vast collections of sequence data were compiled for RNAs, cDNAs, and genomic structures, revealing the base sequences for a number of RNAs. As a result of this work.

(a) Unidentified proteins have been predicted to number 25,243; whereas the known protein number is 15,337.

(b) A majority of mRNA species (95%) mature through alternative splicing mechanisms.

(c) Disease genes are estimated to be 2,577 in number.

(d) Point mutations are 31,250 in number; half of disease-causing mutations are attributed to aberrant splicing (disruption of splicing codes) whereas other forms of mutation include disruption of the genetic code.

(e) Disruption of splicing code occurs at the splice site and enhancer/silencer sites of exonic and intronic sequences.

(f) Pathogenic sequences that occur as a result of splice code mutations (transition and transversion) cause aberrant modifications of a variety of RNAs [68, 69].

Recently, evidence has been accumulating that suggests a need to revise earlier estimates of the number of transcriptional products arising from the genomic information. Paradoxical findings were obtained that contradicted earlier and more conservative estimates of the proteasomes size (50,000), in fact, the cellular transcripts are 2-3 times higher than estimated earlier [50, 51]. Also, 50% of the transcripts were comprised of noncoding RNA, some of which are polyadenylated. This paradoxical manifestation has led to the second generation of genomic work, strictly based on RNA characterization. It is worth emphasizing that this has become the second genomic frontier where a reevaluation of the first genomic work is necessary. The present task is more daunting than the "The first Generation Genome Project." The task at hand is to resequence the genome and then categorize and catalogue the ncRNA species by utilizing all available sequence means, including direct sequencing and DNA microarray techniques.

The next step is to construct secondary structures according to enzyme susceptibility and computer-aided base pairing. Interacting proteins will need to be defined by biochemical, NMR, X-ray, and cryo-EM methods.

5. Reagent and Procedures Required for Sequencing

5.1. RNA-Specific Cleavage Reactions (2′-OH Required Reaction)

(1) Mild alkaline hydrolysis (0.3 N KOH) produces 3′ monophosphorylated nucleotides.

(2) T_1 RNase cleaves phosphodiester bonds after G base producing 3′ GMP at the 3′ ends.

(3) RNase A cleaves phosphodiester bonds after pyrimidines (U and C) producing 3′ phosphates at 3′ ends.

(4) T_2 RNase cleaves all phosphodiester bonds with a preference for A residues, producing 3′ monophosphates.

(5) U_2 RNase cleaves phosphodiester bonds after A base, producing 3′ monophosphates.

The mechanism catalyzed by alkaline hydrolysis, RNase A, T_1 RNase, T_2 RNase and U_2 RNase involves a $S_N2(p)$ mechanism attacking 2′-hydroxyl groups on the adjacent internucleotidic phosphodiester bond to displace the 5′-hydroxyl group of the neighboring nucleotides and generate a 2′, 3′-cyclic nucleotide intermediate. A subsequent hydrolysis of the 2′, 3′-cyclic nucleotide yields a final product, a 3′ mononucleotide (Figure 6).

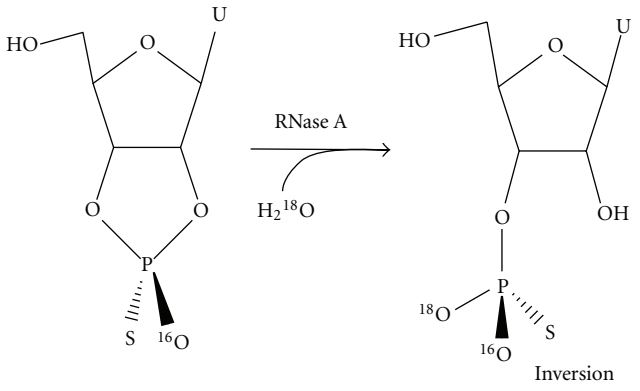
5.2. The Enzymes Cleaving All Phosphodiester Bonds Including 2′-O-Methylated Ribose

(1) P_1 RNase: the enzymatic digestion by P_1 RNase cleaves all phosphodiester bonds (except pyrophosphate linkages), producing 5′ monophosphorylated nucleotides.

(2) The enzymes acting from the ends for sequencing fragments

(a) Snake venom phosphodiesterase (phosphodiesterase I) cleaves phosphodiester bonds, as well as pyrophosphate bonds producing 5′ monophosphorylated nucleotides. It cleaves single-stranded RNA or DNA from the 3′ end in a progressive manner.

(b) Spleen phosphodiesterase (phosphodiesterase II) produces 3′ monophosphorylated nucleotides cleaving from nonphosphorylated 5′ ends of single-stranded RNA or DNA.



Stereochemistry of the reaction catalyzed by RNase A

FIGURE 6: Stereochemistry of the reaction catalyzed by RNase A [10]. The intermediary 2′,3′ cyclic nucleotide (cNp or cNMP) is hydrolyzed to a 3′ phosphorylated mononucleotide. Other 2′-OH requiring enzymatic and alkaline hydrolysis may go through the same path.

5.3. Other Enzymes Utilized for Sequencing

(1) Alkaline phosphatase removes phosphate from 3′ and 5′ ribose moieties.

(2) Pyrophosphatase will only cleave pyrophosphate linkages. There are pyrophosphatases from tobacco and potato as well as from *Crotalus adamanteus* venom type II.

Using varying combinations of fragmentation methods, it becomes possible to obtain fragments that range in size from nucleosides to very large fragments.

5.4. Chemical Modifications Used for Sequencing

5.4.1. CMCT Reaction. Originally reported by Gilham [73], the adduct formation of uridine and guanosine components of RNA with CMCT made uridine residues resistant to RNase A. In addition it has been shown that CMCT reacts with pseudouridine and to a lesser extent with inosine. This reaction takes place on Ψ(N1,N3), U(N3), G(N1), and I(N1), and cold dilute ammonia removes the adducts from Ψ(N1) and hot concentrated ammonia removes remaining adducts from Ψ(N3) [74, 75]. These properties have been used to block RNase A digestion at U but not at C as well as to differentiate U from Ψ (Figure 7) [11].

Direct chemical methods for sequencing RNA using dimethyl sulfate, diethyl pyrocarbonate, and hydrazine followed by aniline-β-elimination have been successfully utilized in 5S RNA and 5.8S RNA sequence analysis [60].

5.4.2. DMS (Dimethylsulfate). This has been used to identify secondary structures as well as for the synthesis of standard $m_3^{2,2,7}$G. The properties of DMS modifying adenosine (N1) and cytosine (N3) make modified nucleotides unable to base-pair. For this reason RT-PCR stops one nucleotide before the modified nucleotide enabling the location of a modified nucleotide as well as differentiating the single-stranded from double-stranded regions of RNA. DMS also has been used for synthesis of $m_3^{2,2,7}$G from N2,N2-dimethylguanosine. For this synthesis, the reaction has been carried out by the methods of Saponara and Enger [76]. Twenty milligrams of N2,N2-dimethylguanosine were suspended in 400 μL of dimethylacetamide containing 10 μL dimethylsulfate. The mixture was shaken for 15 hours at room temperature and then centrifuged to remove insoluble products. The supernatant was adjusted to pH 8.0 with concentrated ammonia and then placed on a phosphocellulose column (1 × 50 cm) at pH 7.0 (0.001 M ammonium acetate). A linear gradient of 0.001–0.3 M ammonium acetate was used to elute the samples. One major peak of the product ($m_3^{2,2,7}$ trimethylguanosine) was found between two minor peaks (corresponding to N2,N2-dimethylguanosine and 7-methylguanosine). The product was lyophilized and identified as $m_3^{2,2,7}$G by mass spectrometry [12]. The summary of reagent and procedures required for sequencing is provided in Table 7.

The nucleotides or nucleosides obtained can be separated by column chromatography, paper electrophoresis or thin layer chromatography to determine the number of G, A. U, C and modified residues in the fragments or in the molecule. These 4 bases have specific UV spectra and chemical reactivity to identify the nature of the bases in comparison with known standards. The unusual nucleoside, trimethylguanosine, has its specific UV absorption spectra (Figure 8) and mass spectrometric characteristics (Figure 9).

6. The Major snRNA Sequenced

The first nuclear small RNA sequenced was 4.5S RNAI [77] shown in Figure 37. This RNA contains the RNA polymerase III promoter box A and box B like motifs and shows interesting enhancer motif elements resembling the Alu element transcript. The RNA polymerase III promoter areas are underlined and the first nucleotide of the enhancer motif is marked by colored letters. The red color is SF2/ASF (4 motifs), blue color is SC35 (3 motifs), green color is SRp40 (6 motifs), and yellow color is SRp55 (1 motif) (Figure 10(a)). It also exhibits 3′-splice sites marked by [AG] as well as branch sites with the highest score marked by {CACCUAU} (Figure 10(b)). The ESE (exonic splice enhancer), splice sites (Figure 10(c)), and branch sites were examined by ESEfinder 3.0 [13].

In comparison with known Alu elements in the FMR1 gene, the resemblance of 4.5S RNA I in ESE, 5′SS, BS, and 3′SS distribution (Table 8) suggests that 4.5S RNA I is more likely derived from an Alu gene expressed in Novikoff hepatoma cells.

The Alu element has been shown to have many different functions in transcription, splicing, exonization [78], gene insertions (transposons), and DNA replication. It is interesting to observe that the (+) oriented Alu has more

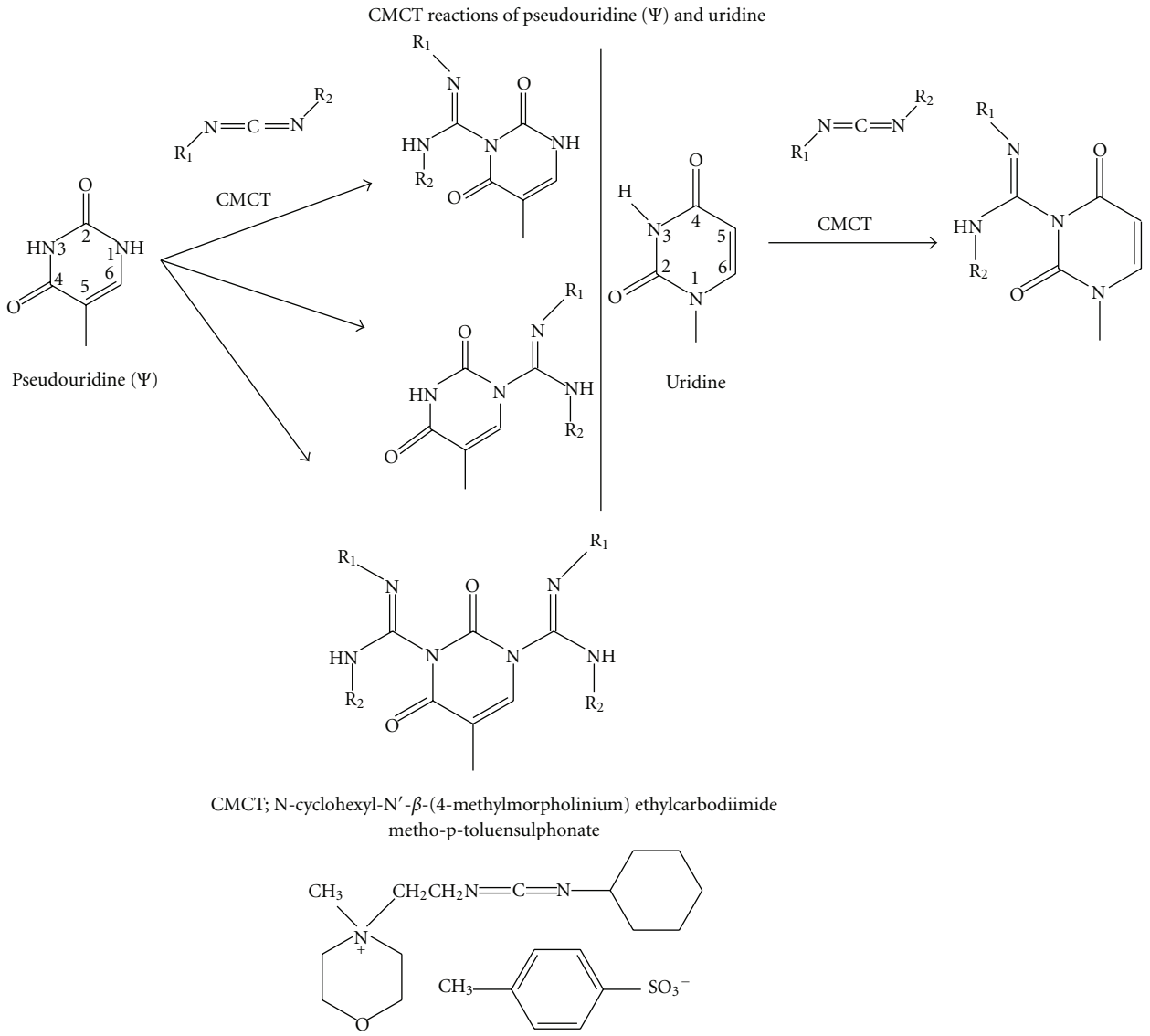


Figure 7: The CMCT reaction of pseudouridine (Ψ) and uridine, and the structure of CMCT [11]. Adducts formed with CMCT on Ψ and U are shown. This adduct formation prevents the cleavage by RNase A at U but not at C. The mild alkaline treatment of reaction products destroys the U but not the Ψ. These differences were utilized to locate the position of Ψ by reverse transcriptase.

5′ splice sites and the (−) oriented Alu has more 3′ splice sites. It may suggest that exonization may occur from the 5′ side of (+) Alu elements and 3′ side of (−) Alu elements. The SRP RNA (7SL RNA) has Alu elements in its sequence [79]. Whether the Alu is derived from 7SL or Alu is exonized to 7SL is not clear. Subsequently, other snRNAs have been sequenced.

The sequences of the capped snRNAs are described in Figure 11. The pivotal sequences needed for functions are marked by colors.

In the course of any sequence work, there are always challenges in resolving unknown structures at the 5′ end portions which contain the 5′-cap structure and various modified nucleotides. The experimental steps required to discern this complicated region are described.

7. Nucleotide Composition and Modified Nucleotides in snRNAs

The compositional analyses were carried out by UV analysis as well as isotope labeling analysis. For example, UV analysis required ~10 mg of U2 RNA.

7.1. RNA Terminal Labeling with [^{3}H]-KBH$_4$. The purified nuclear RNAs were separated by sucrose gradient centrifugation which separates 4–8S RNA, 18S RNA, 28S RNA, 35S RNA, and 45S RNA isolated from nuclei of rat liver, Walker tumor, or Novikoff hepatoma cells. As an initial step for the structural characterization, 3′ end nucleosides were labeled by the procedure of sodium periodate ($NaIO_4$) oxidation and potassium borohydride ([^{3}H]-KBH_4) reduction.

Table 7: Reagents and procedures required for sequencing.

Detection of external modification	Oligo-dT column Borate column Antibodies against m^7G, $m_3^{2,2,7}G$, and others
Labeling	Prelabeling (in vivo) with [^{32}P]-phosphate Postlabeling (in vitro) 5′ [^{32}P] labeling 3′ [^{32}P] labeling 3′ and 5′ [^{3}H]-derivative labeling
Enzymes	
Endonucleases	RNase T1(Gp↓N), RNase A(Up↓N or Cp↓N), RNase U2(Ap↓N), PhysI(C-resistance), RNase T2(Np↓N), RNase P1(pNm↓pN, pN↓pN), Debranching enzyme N (2′-↓5′pNp / 3′-p-Np)
Exonuclease	Spleen phosphodiesterase (Np↓Np↓Np) → Snake venom phosphodiesterase (N↓pN↓pNp and N5′p↓p↓p5′N) ← → ← Pyrophosphatase (-p↓p-)
Phosphatase	Phosphatase (p↓N and N↓p)
Chemicals	
CMCT	React with U>Ψ>G>I on ssRNA
DMS	Methylate G>A>C on ssRNA
Diethyl pyrocarbonate	React with A>G
Hydrazine	React with U>>C>T
Limited hydrolysis	Formamide Hot water (80°C)
Fractionation	2D acrylamide gel electrophoresis 2D homochromatography 2D electrophoresis 2D paper chromatography 2D thin layer chromatography

The reaction was carried out in 0.1 M sodium acetate buffer at pH 5 with freshly prepared $NaIO_4$ in the dark for 1 hour and precipitated the RNA with ethanol. The RNA was redissolved in the same buffer and treated with ethylene glycol to destroy excess $NaIO_4$. The RNA was precipitated with ethanol and redissolved in 0.1 M sodium phosphate buffer, pH 7.7, and treated with radioactive [^{3}H]-KBH_4 [38]. These reaction products would have tritium labeling in cis-alcohols from cis-aldehyde oxidation products of the 2′ and 3′ hydroxyls of ribose, assuming all 3′ ends of RNA have accessible 2′ and 3′ OH groups (Figure 12).

The labeled 4–8S RNAs were separated by preparative polyacrylamide gel electrophoresis (Figure 13) and DEAE-Sephadex column chromatography (Figure 14) to purify individual snRNAs (U1, U2, U3, 4.5S RNA I, II, and III, 5S RNA I, II, and III).

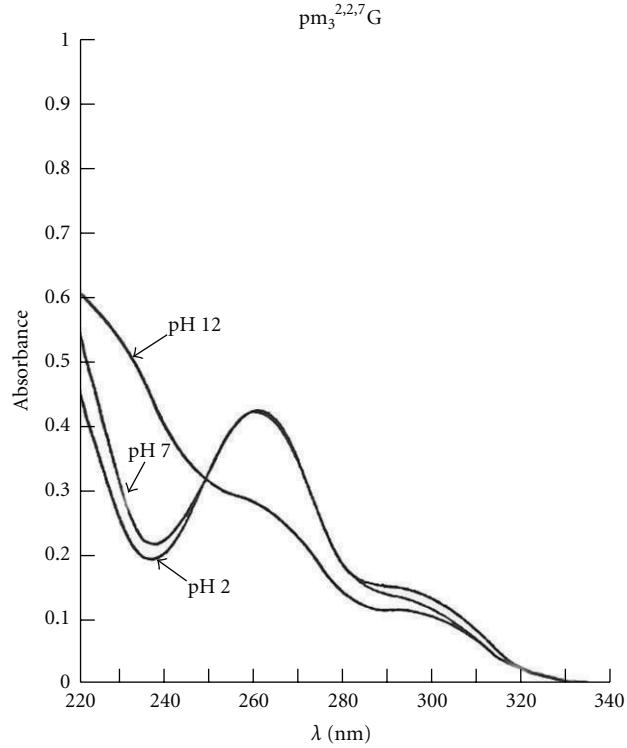


Figure 8: The UV spectra of $pm_3^{2,2,7}G$ [12]. The ultraviolet absorption spectra were recorded on a Cary 14 spectrometer immediately after addition of compound to solutions at pH 2, 7, and 12.

Alkaline hydrolysis of these RNAs produced 3′ end nucleoside trialcohol derivatives (Table 9) which were subsequently identified by paper chromatography.

The RNA that appeared to be pure for sequencing was 4.5S RNA I which had 87.4% U at the 3′ terminus and only 6.5% unknown radioactivity at the origin. Unexpectedly, U1, U2, U3, 4.5S RNA II, and some of 5S RNA (5S RNA III/U5) had ~50% labeling in alkaline-resistant fragments that did not move as nucleoside derivatives. The 4.5S RNA III was not labeled by this procedure suggesting a blocked 3′ end (Figure 14). The U1, U2, and U3 RNAs were labeled with tritium, digested with RNase A, and separated on a DEAE-Sephadex column (Figure 15).

The oligonucleotides were digested with T_1 RNase and rechromatographed, and only the U3 oligonucleotide was shortened by one nucleotide, indicating the presence of one G adjacent to RNase A susceptible pyrimidine [80]. In the course of sequencing U1, U2, U3 RNAs, it was found that the oligonucleotides with $m_3^{2,2,7}G$ was coming from the 5′ end segments. The only way 2′3′ hydroxyls could be at 5′ end was 5′-5′ pyrophosphate linkage to the rest of the RNA molecules [36]. The RNase A and T_1RNase resistant oligonucleotides were digested with various enzyme combinations including snake venom phosphodiesterase, alkaline phosphatase, P_1 RNase, T_2 RNase, and U_2 RNase into nucleosides. The component nucleosides were identified by mass spectrometry, U.V. spectroscopy, HPLC (high pressure liquid chromatography), paper chromatography, and thin layer chromatography. [12, 16, 37, 58].

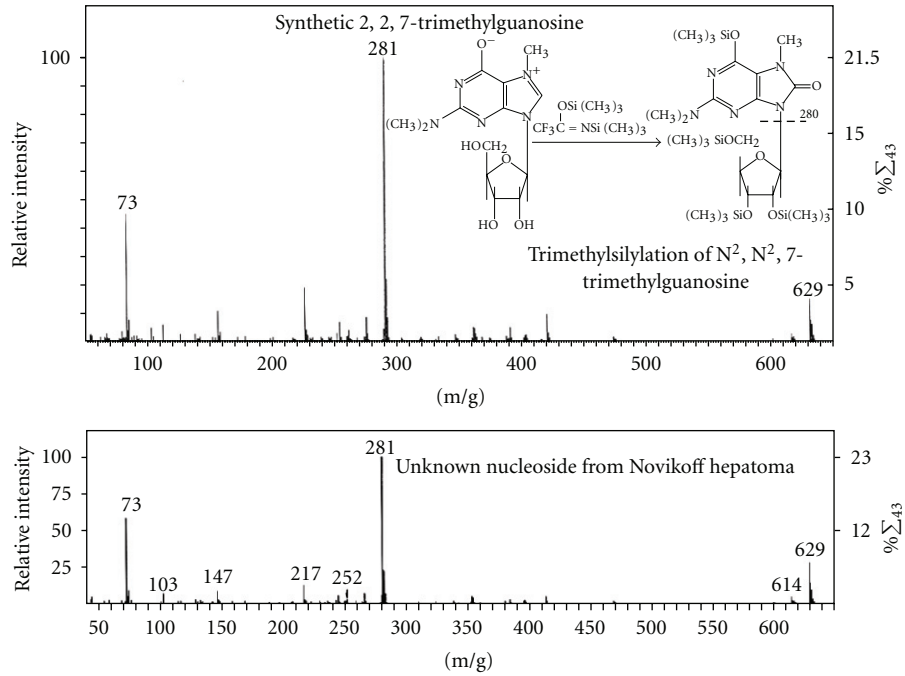


Figure 9: The mass spectra of trimethylguanosine [12]. The synthetic $m_3^{2,2,7}G$ and unknown nucleoside from U2 RNA were trimethylsilylated and subjected to LKB 9000 gas chromatograph-mass spectrometer. The mass spectrum of the unknown nucleoside from U2 RNA was identical to synthetic $m_3^{2,2,7}G$.

Table 8: Distribution of ESE, 5′ splice sites, branch sites, and 3′ splice sites. The ESE, 5′ splice site, branch site, and 3′ splice site in 4.5S RNA I and Alu elements in FMR1 gene transcript are screened by ESE finder (version 3) [13]. For this comparison, the number of motifs is calculated per 100 nucleotides. The motif patterns in Alu elements are all very much alike and the 4.5S RNA I resembles them. A difference is found in that 5′ splice sites in (+) Alu are more than in (−) Alu and 3′ splice sits are more in (−) Alu than in (+) Alu.

	SF2/ASF	SC35	SRp40	SRp55	Total	5′SS	BS	3′SS
Novikoff 4.5S RNA I (96 nt)	3.65	3.13	6.25	1.04	**14.07**	0	10.4	2.08
Human FMR1 Alu1(+) (252 nt)	4.96	7.54	6.35	1.59	20.44	3.97	9.13	5.56
Human FMR1 Alu4(+) (295 nt)	4.41	4.41	1.69	2.03	12.54	2.37	10.9	2.37
Human FMR1 Alu5(+) (246 nt)	5.69	5.28	4.47	1.22	16.66	4.47	8.54	2.44
Human FMR1 Alu7(+) (290 nt)	6.21	2.41	3.10	0.34	12.06	3.79	10.3	2.07
Human FRM1 Alu8(+) (288 nt)	5.21	3.47	3.82	0.69	13.19	3.47	8.68	2.43
(+) Alu Av.	5.30	4.62	3.89	1.17	**14.98**	**3.61**	9.51	2.97
Human FMR1 Alu2(−) (298 nt)	3.69	6.04	5.03	1.34	16.10	0.67	14.1	4.70
Human FMR1 Alu3(−) (285 nt)	4.92	7.02	5.61	2.11	19.66	3.16	13.7	4.56
Human FMR1 Alu6(−) (290 nt)	4.31	5.52	4.48	1.38	15.69	2.76	13.5	4.48
(−) Alu Av.	4.31	6.19	5.04	1.61	**17.15**	2.20	13.8	**4.58**
Human FMR1 Total Alu(+&−) Av.	4.93	5.21	4.32	1.34	**15.80**	3.08	11.1	3.58

7.2. Tritium Labeling of Nucleosides. The purified RNAs were digested with RNase A, snake venom phosphodiesterase, and alkaline phosphatase at pH 8.0, 37°C for 6 hours into nucleosides. The digest was treated with a 2X molar excess of $NaIO_4$ and labeled with $[^3H]$-KBH_4 at pH 6 for 2 hours in the dark to produce trialcohol derivatives of nucleosides. All nucleosides with base modifications, except 2′-O-ribose modified, were labeled with tritium. The tritium-labeled trialcohol derivatives were separated by two-dimensional TLC (thin layer chromatography) on cellulose thin layers (Figure 5) [81]. The first dimension used a solvent of acetonitrile, ethylacetate, n-butanol, isopropanol, 6 N aqueous ammonia (7 : 2 : 1 : 1 : 2.7); the second dimension used a solvent of tert-amyl alcohol, methylethylketone, acetonitrile, ethylacetate, water, formic acid (sp.gr. 1.2) (4 : 2 : 1.5 : 2 : 1.5 : 0.18) [81, 82].

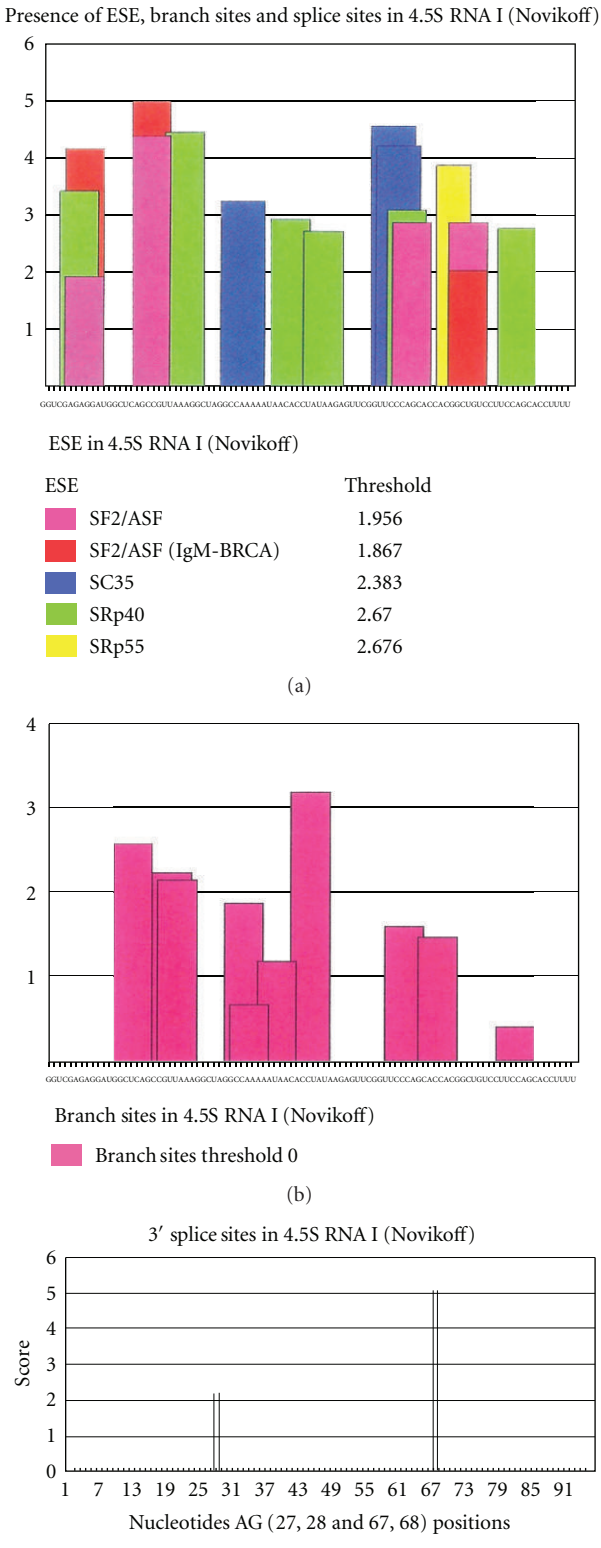


Figure 10: ESE motifs in 4.5S RNA. The sequences of 4.5S RNA I from Novikoff hepatoma cell nuclei were screened by ESE finder 3 [13] for ESE, 5′ splice sites, branch sites, and 3′ splice sites. The default threshold value was used. There were 4 SF2/ASF sites, 3 SC35 sites, 6 SRp40 sites, 1 SRp55 sites, 10 branch sites, and 2 3′ splice sites. These numbers resemble the number identified in Alu elements of human FMR1 transcript (Table 8).

Table 9: The snRNA 2′ and 3′-OH labeling by $NaIO_4$ oxidation and [^{3}H]-KBH_4 reduction [38]. The total 4–7S RNA from rat Novikoff hepatoma cell nuclei was labeled with [^{3}H] by oxidation with $NaIO_4$ followed by [^{3}H]-KBH_4 reduction (Figure 12). Individual RNA species were purified by gel electrophoresis (Figure 13). The RNA samples were hydrolyzed with 0.3 N KOH, and hydrolysates were chromatographed on whatman 3MM paper according to de Wachter and Fiers [55]. The radioactivities at the origin (22% for 5S RNA, 54.1% for U1 RNA, 49.7% for U2 RNA, and 50.6% for U3 RNA) represent % of total radioactivity applied and they represent the 5′ end labeling which was later elucidated by many enzymatic methods described in the text. The radioactivities moved by chromatography with standard nucleoside derivatives are the % of total in nucleosides derivatives. The A′ U′ G′ C′ represent trialcohol derivatives of nucleosides.

RNA Species	Radioactivity at origin (5′)	3′ Nucleoside derivatives			
		A′	U′	G′	C′
	%	%	%	%	%
4S RNA	10.9	89.0	3.2	3.9	3.8
4.5S RNA	15.8	11.2	79.7	4.0	5.1
4.5S RNA I	6.5	6.1	87.4	4.7	1.8
4.5S RNA II	30.9	13.1	80.2	4.7	2.4
5S RNA	22.0	11.4	75.5	6.0	7.0
U1 RNA	54.1	6.0	13.4	77.7	3.0
U2 RNA	49.7	61.5	6.0	4.3	28.2
U3 RNA	50.6	53.8	22.6	9.8	13.7

7.3. [^{32}P] Labeling of RNA. The Novikoff hepatoma cells were transplanted intraperitoneally into male albino rats of the Holtzman strain weighing 200–250 g, obtained from Cheek Jones Company (Houston, Tex). After 5-6 days, the cells were harvested and washed with NKM solution (0.13 M NaCl, 0.005 M KCl, and 0.008 M $MgCl_2$). Twenty milliliter (packed volume) of cells was incubated with 500 mCi of [^{32}P]-orthophosphate in 1 liter of medium (phosphate free modified Eagle's minimal essential medium) for 9–16 hours [83]. Nuclear RNA was purified by sucrose gradient centrifugation, gel electrophoresis, and column chromatography [38]. The purified RNA was hydrolyzed with 0.3 N KOH, and alkaline-resistant oligonucleotides were separated on DEAE-Sephadex. The alkaline resistant dinucleotides were collected, treated with alkaline phosphatase, and identified by two-dimensional chromatography (Figure 16).

The summary of modified nucleotides is in Table 1 [84].

8. Structural Determination of 5′ Oligonucleotides

The structures of the 5′ ends of U1 RNA, U2 RNA, U3 RNA, and 5S RNA III (U5) are determined by the characteristics of chemical reactions and enzymatic susceptibilities (Figure 17).

8.1. U1 RNA 5′ End Oligonucleotide. The U1 RNA labeled with [^{3}H] by $NaIO_4$ and [^{3}H]-KBH_4, digested with RNase

(1) Human HeLa U1A RNA (2) Rat Novikoff U1 RNA (Involved pre-mRNA splicing)

(1) $m_3^{2,2,7}$GpppAmUmACΨΨACCU GGCAGGGGAG AUACCAUGAU
(2) $m_3^{2,2,7}$GpppAmUmACΨΨACCU GGCAGGGAGA UACCAUGAUC

(1) CACGAAGGUG GUUUUCCCAG GGCGAGGCUU AUCCAUUGCAm CUCCGGAUGU
(2) ACGAAGGUGU UUUCUCUCCA GGGCGUGGUC UAUCCAUUGA GGCGCAmCUCC

(1) GCUGACCCCU GCGAUUUCCC CAAAUGUGGG AAACUCGACU GCAUAAUUUG
(2) GUGGAUGCUG ACCCCUGCGA UUUCCUCCAA AUGCGGGAAA CUCGACUGCA

(1) UGGUAGUGGG GGACUGCGUU CGCGCUUUCC CCUG-OH (164) [15]
(2) UAAUUUGUGG UAGUGCGGGG GACUGUUCGC GCUCCUCUCG-OH (170) [16]
Pre-mRNA 5′ splice site binding region Sm protein binding region

(3) Human HeLa U2 RNA (4) Rat Novikoff U2 RNA (Involved in pre-mRNA splicing)

(3) $m_3^{2,2,7}$GpppAmUmCGCUUCUC GmGmCCUUUUGmG CUAAGmAUCAAm
(4) $m_3^{2,2,7}$GpppAmUmCGCGmGmCUC CΨΨCΨUUUUGm GCUAAGmAUCA

(3) GUGUAGUAUCm ΨGUUCUUAUC AGUUUAAUAU CmUGAUACGUC
(4) m^6AmGUGΨAGΨAΨ CmΨGΨΨUCUmAU CAGUΨUUAAΨ AUCmUUCGAUA

(3) CUCUAUCCGA GGACAAUAUA UUAAAUGGAU UUUUGGAGCA GGGAGAUGGA
(4) CGUCCUCUAU CCGAGGACAA UAΨUAΨUAAA UGGAUUUUUG GAACUAGGAG

(3) AUAGGAGCUU GCUCCGUCCA CUCCACGCAU CGACCUGGUA UUGCAGUACC
(4) UUGGAAUAGG AGCUUGCUCC GUCCACCUCA CGCAUCGACC UGGUAUUGCG

(3) UCCAGGAACG GUGCACCA-OH (188) [17]
(4) CAGUACCCUC AGGAACGGUG CACCA-OH (195) [18]
Branch point binding region Sm protein binding region

(5) Human Placenta U3 RNA (6) Rat Novikoff U3B RNA (Involved in rRNA production)

(5) $m_3^{2,2,7}$GpppAmAmGACUAUAC UUUCAGGGAU CAUUUCUAUA GUGUGUUACU
(6) $m_3^{2,2,7}$GpppAmA(m)AGCUAUAC UUUCAGGGAU CAUUUCUAUA GUUCGUUACU

(5) AGAGAAGUUU CUCUGAACGU GUAGAGCACC GAAAACCACG AGGAAGAGAG
(6) AGAGAAGUUU CUCUGACUGU GUAGAGCACC CGAAACCACG AGGACGAGAC

(5) GUAGCGUUUU CUCCUGAGCG UGAAGCCGGC UUUCUGGCGU UGCUUGGCUG
(6) AUAGCGUCCC CUCCUGAGCG UGAAGCCGGC UCUAGGUGCU GCUUCUGUGC

(5) CAACUGCCGU CAGCCAUUGA UGAUCGUUCU UCUCUCCGUA UUGGGGAGUG
(6) AGCUGCCUCU UGCCAUUGAU GAUCGUUCUU CUCUCCUUCG GGAGGGUAAG

(5) AGAGGGAGAG AACGCGGUCU GAGUGGU-OH (217) [19,20]
(6) AGGGAGGGAA CGCAGUCUGA GUGGA-OH (215) [21]
Box A Box B Box C Box D

(7) Human HeLa U4A RNA (8) Rat Novikoff U4 RNA (Involved in pre-mRNA splicing)

(7) $m_3^{2,2,7}$GpppAmGmCΨUUGCmGC AGUGGCAGUA UCGUAGCCAA UGAGGUUUAU
(8) $m_3^{2,2,7}$GpppAmACΨUUGCAG UGGCAGUAUC GUAGCCAAUG AGGUUUAUCC

(7) CCGAGGCGCG AUUAUUGCUA AUUGAmAAACU UΨUCCCAAΨA CCCCGCCGUG
(8) GAGGCGCGAU UAUUGCUAAU UGAmAAACUUU UCCCAAΨACC CCGCCGUGAC

(7) ACGACUUGCA AUAUAGUCGG CAUUGGCAAU UUUUGACAGU CUCUACGGAG
(8) GACUUGAm^{6}AAU AUAGUCGGCA UUGGCAAUUU UUGUCUCUAC GGAGACUG

(7) ACUG(G)-OH (144-145)[22,23]
(8) (138) [24]

(a)

Figure 11: Continued.

U6 RNA binding region 15.5 K protein & hPrp31 binding region [23]

(9) Human HeLa U5A RNA (10) Rat Novikoff U5A RNA(Involved in pre-mRNA splicing)

(9) $m_3^{2,2,7}$GpppAmUmACUCUGGU UUCUCUUCAG AUCGCAUAAA UCUUUCGmCCU
(10) $m_3^{2,2,7}$GpppAmUmACUCUGGU UUCUCUUCAG AUCGUAUAAA UCUUUCGmCCU

(9) UmUΨACmΨAAAGAUΨUCCGUGG AGAGGAACAA CUCUGAGUCU UAACCCAAUU
(10) UmUUAUmCAAAGAUΨUCCGUGGAGAGGAACAACUCUGAGUCU UAAACCAAUU

(9) UUUUGAGCCU UGCCUUGGCA AGGCUA-OH (116) [25]
(10) UUUUGAGGCC UUGUCUUGAC AAGGCUA-OH (117) [26]
mRNA binding region in spliceosome

(11) Human Liver U6 RNA (12) Rat Novikoff U6 RNA(Involved in pre-mRNA splicing)

(11) mpppGUGCUCGCUU CGGCAGCACA UAUACUAAAA UUGGAACGAU
(12) mpppGUGCCUGCUU CGGCAGCACA UAUACUAAAA ΨUGGAACGAΨ

(11) ACm^{6}AGAGAmAGAUUAmGmCAUGGCmCCmCmUGCGCAAm Gm^{2}GAUGACmACG
(12) ACm^6 AGAGAmAGAUUAmGmCAUGGCm CCmCmUGCGCAAm Gm^2 GAUGACmACG

(11) CAAAUUCGUG AAGCGUUCCA UAUUUU(U)-OH (106-107) [27]
(12) CAAAUΨCGUG AAGCGUUCCA UAUUUU(U)-OH (106-107) [28]
U4 RNA binding region

(13) Human HeLa U7 RNA (Involved in Histone mRNA processing)

(13) $m_3^{2,2,7}$GpppNNGUGUUACA GCUCUUUUAG AAUUUGUCUA GUAGGCUUUC

(13) UGGCUUUUCA CCGGAAAGCC CCU-OH (63) [29]
Histone pre-mRNA 3′ end binding region Sm & LSm protein binding region

(14) Human HeLa U8 RNA (Involved in rRNA processing)

(14) $m_3^{2,2,7}$GpppAUCGUCAGGU GGGAUAAUCC UUACCUGUUC CUCCUCCGGA

(14) GGGCAGAUUA GAACAUGAUG AUUGGAGAUG CAUGAAACGU GAUUAACGUC

(14) UCUGCGUAAU CAGGUCUUGC AACAUCCUGA UUGCUCCUAU CUGAUU-OH
(136) [30]
Box C Box D

(15) Human U13 RNA (Involved in rRNA processing)

(15) $m_3^{2,2,7}$GpppGAUCCUUUUG UAGUUCAUGA GCGUGAUGAU UGGGUGUUCA

(15) UACGCUUGUG UGAGAUGUGC CACCCUUGAA CCUUGUUACG ACGUGGGCAC

(15) AUUAUUCGUC UGAUC-OH (105)[30]
Box C Box D 18S rRNA binding region

(b)

Figure 11: Sequences of major snRNAs (see [15–30]). The sequences of major snRNAs from human and rat involved in splicing and processing are aligned for comparison. The sequence elements in major spliceosomal snRNAs and processosomal snoRNAs are highlighted in the sequences. Those are the pivotal motifs for the function. The numbers in parenthesis are the chain length of the RNAs.

A, showed enzyme-resistant oligonucleotide eluting close to the pentanucleotide region on a DEAE column (Figure 15). The 5′ oligonucleotide was analyzed by UV, [^{3}H], and [^{32}P] methods.

8.1.1. The UV Analysis. The 5′ oligonucleotides from U1 RNA, obtained by RNase A and RNase T1, were digested with snake venom phosphodiesterase and alkaline phosphatase. The nucleosides produced were separated on HPLC (high pressure liquid chromatography) [strongly basic cation exchange (quaternary amine)]. As shown in Figure 18, the amount of nucleoside ratio was 1.0, 1.2, 1.2, 0.7, and 0.9 for Am, A, Um, $m_3^{2,2,7}$G, and C, respectively, for U1 5′ oligonucleotide.

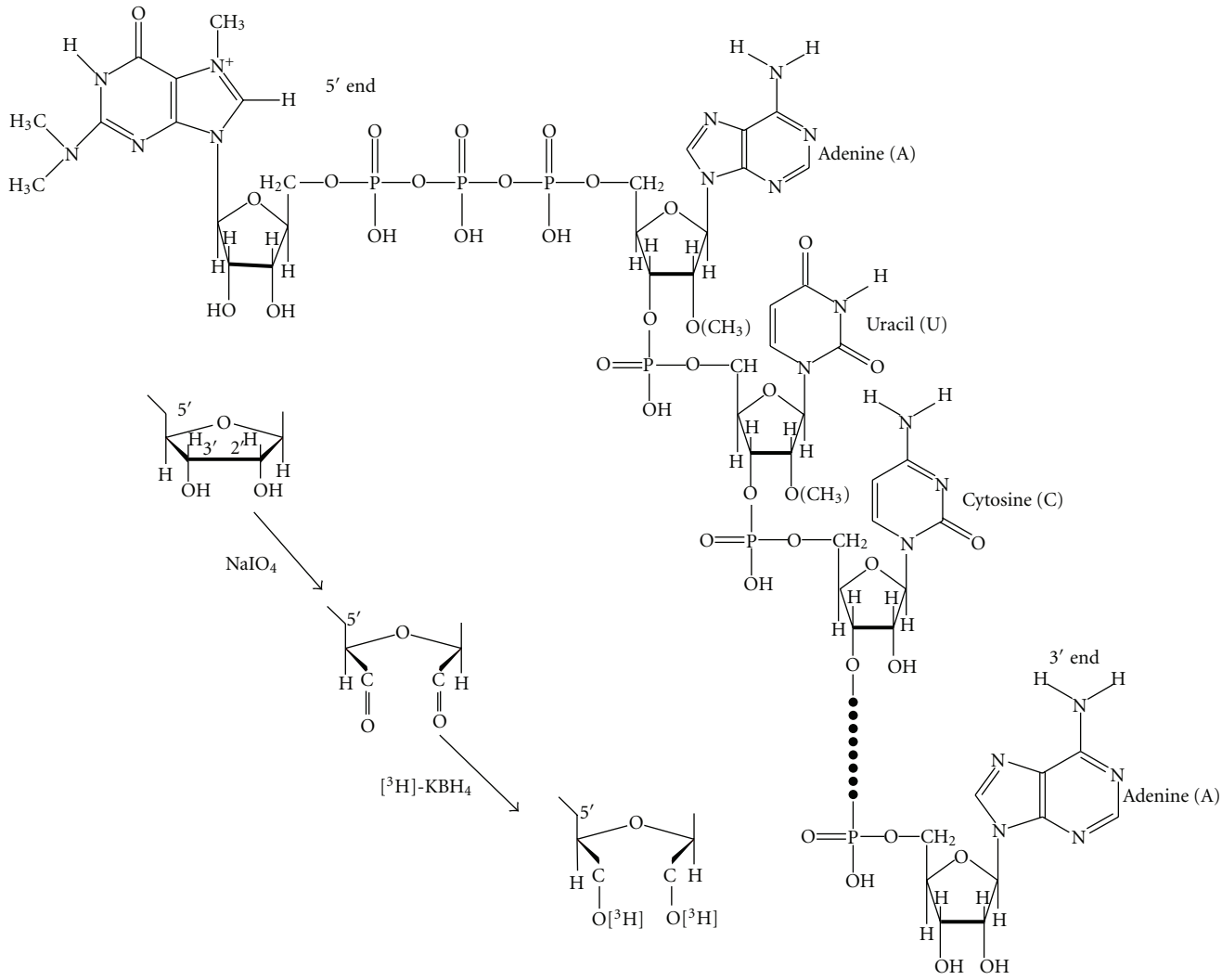


Figure 12: RNA 5′ and 3′ end labeling with [^{3}H] by treatment with $NaIO_4$ and [^{3}H]-KBH_4. The 4–7S RNA from Novikoff hepatoma cell nuclei was labeled to detect the presence of free 2′-OH and 3′-OH by $NaIO_4$ oxidation followed by [^{3}H]-KBH_4reduction. The reaction occurred at both ends of RNA (5′ end and 3′ end).

8.1.2. The [^{3}H] Method. The [^{3}H]-labeled U1 RNA 5′ oligonucleotide, following digestion with snake venom phosphodiesterase and alkaline phosphatase, was separated by chromatographic methods with standards. Two-dimensional TLC (thin layer chromatography) and paper chromatography demonstrated that the [^{3}H] labeled compound is a trimethylguanosine derivative (Figure 19).

8.1.3. ^{32}P-Labeled 5′ Oligonucleotide from U1 RNA. The ^{32}P-labeled RNA was digested with T_2 and U_2 RNase, and digestion products were separated by two-dimensional electrophoresis. The first dimension was on cellogel at pH 3.5, and the second dimension was on DEAE paper at pH 3.5 (Figure 20).

Spot "a" was eluted and treated with alkaline phosphatase and chromatographed with GMP, GDP, and GTP standards. The ^{32}P-labeled 5′ oligonucleotide was chromatographed in the GTP region on a DEAE-Sephadex column (Figure 21).

The oligonucleotide peak from the GTP region was digested with snake venom phosphodiesterase and separated by electrophoresis in the first dimension followed by chromatography on second dimension (Figure 22).

The ^{32}P activity ratio was 1.00, 1.11, 1.25, 0.53, and 1.14 for $pm_3^{2,2,7}G$, pAm, pUm, pA, and Pi, respectively. The peak from the GTP region in Figure 21 digested with RNase P1 produced pUm, pA (peak a in Figure 23), and cap core $m_3^{2,2,7}$GpppAm (peak b in Figure 23). Table 10 shows the radioactivity distribution in peaks a and b in Figure 23.

For the analysis of a number of phosphates in cap core (peak b), the cap core was treated with $NaIO_4$ and aniline to remove $m_3^{2,2,7}G$ by β-elimination reaction (Figure 24).

The product was chromatographed on a DEAE column with standard AMP, ADP, and ATP. The product was eluted close to ATP, indicating that it is pppAm. This experiment proved that the 5′ oligonucleotide structure is **$m_3^{2,2,7}$GpppAmpUmpApCp**.

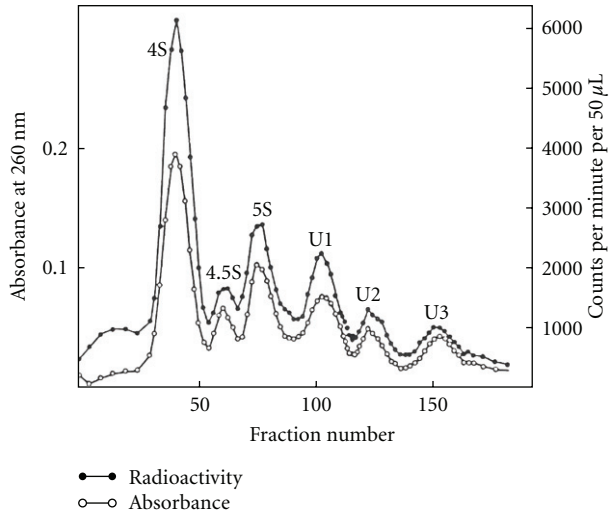


Figure 13: Preparative gel electrophoretic pattern of 4–7S RNA labeled with [^{3}H] [16]. The intact, labeled RNA (as in Figure 12) was subjected to preparative polyacrylamide gel electrophoresis. The UV absorption and radioactivity were measured.

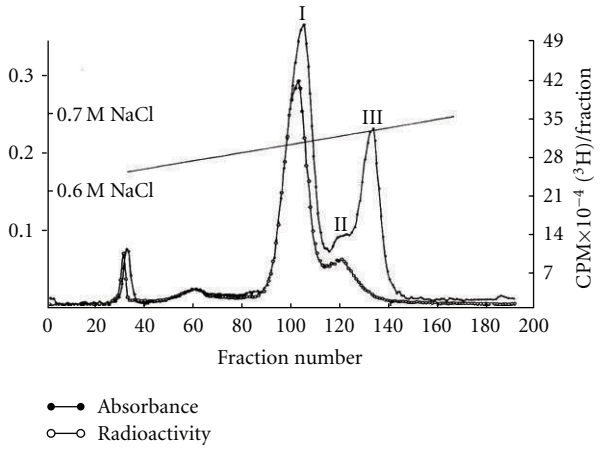


Figure 14: DEAE-Sephadex column chromatography of 4.5S RNA [38]. The 4.5S RNA from preparative gel electrophoresis was collected and chromatographed on a DEAE-Sephadex A-50 column. It was separated into 3 peaks but no radioactivity in 4.5S RNA III was detected, indicating the absence of accessible 2′-OH and 3′-OH in this molecule. The 4.5S RNA II may be the U6 RNA.

8.2. U2 RNA 5′ End Oligonucleotide. The U2 RNA labeled with $NaIO_4$ and [^{3}H]-KBH_4 was digested with RNase A. The labeled oligonucleotide eluted around the tetranucleotide region (Figure 15). The 5′ oligonucleotide was analyzed by UV, [^{3}H], and [^{32}P] methods.

8.2.1. UV Analysis. The 5′ oligonucleotide obtained by complete RNase A digestion was analyzed for its base composition. The purified 5′ oligonucleotide was digested with snake venom phosphodiesterase followed by alkaline phosphatase. The digestion product (nucleosides) was separated by HPLC. The composition was Am, Um, C, and $m_3{}^{2,2,7}$G in a ratio of

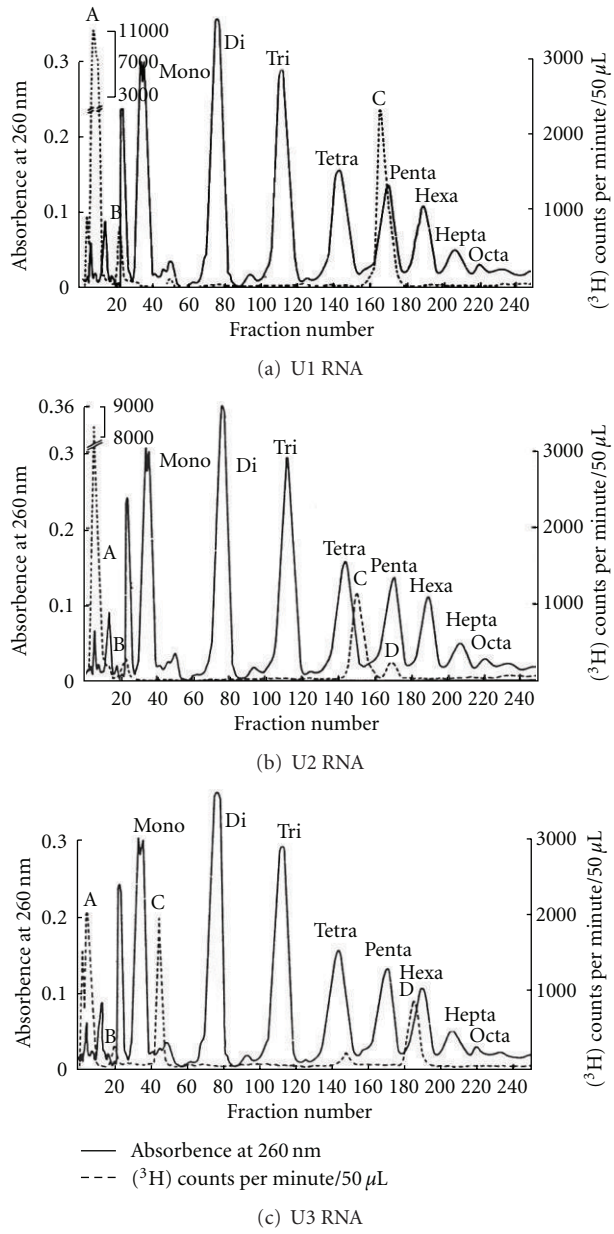
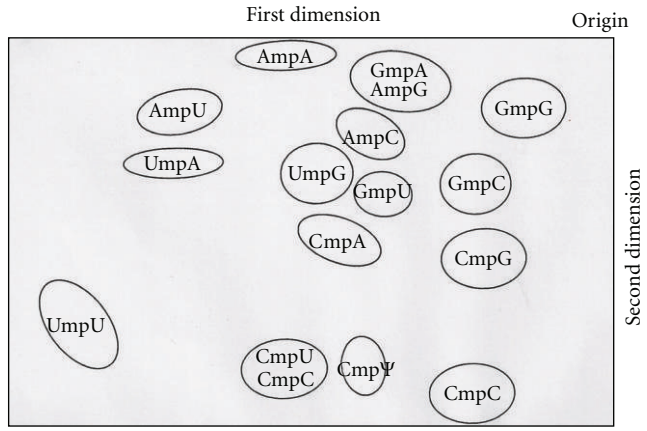
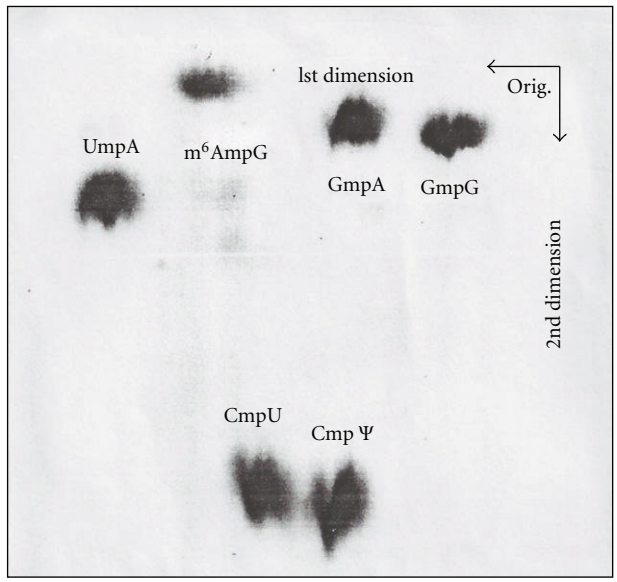


Figure 15: DEAE-Sephadex chromatography of 5′ oligonucleotide [37]. The U1 RNA, U2 RNA, and U3 RNA collected from preparative gel electrophoresis (Figure 13) were digested with RNase A and subjected to DEAE-Sephadex A-25 column chromatography. The radioactive peaks at nucleoside region were coming from 3′ ends and the radioactivities at the regions of penta-, tetra-, and hexanucleotides were from 5′ end labeling. These fragments were treated with T1 RNase and found to be shortened by one nucleotide only in U3 5′ oligonucleotide indicating that G was next to the terminal pyrimidine nucleotide.

1.0, 1.3, 1.1, and 0.96, respectively, (Figure 18) [12]. These nucleosides were also separated by two-dimensional TLC in a borate system. Um and Am migrated through the butanol-boric acid while the $m_3{}^{2,2,7}$G and C, which form complexes with borate, were retarded in the butanol-boric acid phase (Figure 25).



(a)



(b)

Figure 16: Two-dimensional separation of alkali stable dinucleoside monophosphate [18, 39]. (a) Standard dinucleoside monophosphates (NmpN) were separated on Whatman no.1 paper with the solvent systems ethylacetate-1-propanol-water (4 : 1 : 2, v/v) in the first dimension and in the second dimension with 2-propanol-water-concentrated ammonium hydroxide (7 : 2 : 1, V/V). (b) Autoradiograph of two-dimensional separation of alkali-stable dinucleoside monophosphate of U2 RNA from Novikoff hepatoma cell nuclei. The [^{32}P]-labeled U2 RNA was hydrolyzed by 0.3 N NaOH, and the sample was separated on a DEAE-Sephadex column A-25 at pH 7.6. The dinucleotides were collected and treated with alkaline phosphatase. The alkali stable dinucleoside monophosphates were separated on Whatman no. 1 paper and autoradiographed with X-ray film.

The UV spectra of $pm_3^{2,2,7}G$ were typical of a trimethyl G nucleotide (Figure 8). The mass spectrometry of the unknown nucleoside from U2 RNA 5′ fragment was identified as $m_3^{2,2,7}$ trimethylguanosine (Figure 9).

8.2.2. [^{3}H]-Labeled U2 RNA 5′ Oligonucleotide. The purified U2 RNA, labeled with $NaIO_4$ and [^{3}H]-KBH_4 methods, was digested with RNase A and 5′ oligonucleotide purified by DEAE-Sephadex column chromatography (Figure 15). The purified 5′ oligonucleotide was digested with snake venom phosphodiesterase followed by alkaline phosphatase. The nucleosides obtained were separated on two-dimensional TLC [12] and 3MM paper chromatography. The tritium-labeled compound was identified as a trialcohol derivative of $m_3^{2,2,7}G$ (Figure 26).

Table 10: The [^{32}P]cpm radioactivity distribution in nucleotides and cap core produced by RNase P1 digestion of U1 5′ oligonucleotide [16]. The [^{32}P]-labeled U1 5′ oligonucleotide obtained by digestion of U1 RNA with RNase T2, RNase U2, and alkaline phosphatase was treated with P_1 nuclease which cleaves all phosphodiester bonds but not pyrophosphate bonds. The products were separated on a DEAE column (Figure 23). The radioactivity in peak a (mononucleotides pUm, pA) and peak b (cap core $m_3^{2,2,7}$GpppAm) were determined by Packard liquid scintillation spectrometer.

	Peak a (pUm, pA)	Peak b ($m_3^{2,2,7}$GpppAm)
U1 5′ cap	65,280	87,600

8.2.3. [^{32}P]-Labeled U2 RNA 5′ Oligonucleotide. The [^{32}P]-labeled U2 RNA was digested with T_1 RNase or RNase A. Half of each 5′ oligonucleotide was digested with alkaline phosphatase. Oligonucleotides were subsequently digested with snake venom phosphodiesterase, and the resulting 5′ nucleotides were separated first by electrophoresis and second by chromatography (Figure 27). The ratio of [^{32}P] counts is shown in Table 11.

The U2 RNA 5′ oligonucleotide obtained by RNase A was subjected to digestion with pyrophosphatase (*Crotalus adamanteus* venom type II, Sigma). The remaining oligonucleotide did not have $m_3^{2,2,7}G$, indicating that the $m_3^{2,2,7}G$ is linked by pyrophosphate linkage (Figure 28).

From these data the 5′ end oligonucleotide from U2 RNA has been deduced to be **$m_3^{2,2,7}$GpppAmpUmpCpGp**.

8.3. U3 RNA 5′ End Oligonucleotide. The [^{3}H]-labeled U3 RNA was digested with RNase A and or T_1 RNase. The [^{3}H]-labeled 5′ oligonucleotide obtained by RNase A digestion was eluted in the hexanucleotide region (Figure 15). The [^{32}P]-labeled U3 RNA digested with T_2 and U_2 RNA produced 2 spots that were separated by two-dimensional electrophoresis (Figure 29).

8.3.1. UV Analysis. The 5′ oligonucleotide obtained from U3 RNA by digestion with RNase A and T_1 RNase was isolated by column chromatography. The purified 5′ oligonucleotide was digested with snake venom phosphodiesterase and alkaline phosphatase. The nucleosides obtained were subjected to HPLC. The molar ratios of $m_3^{2,2,7}G$, Am, A, and G were 1.0, 1.7, 1.1, and 1.0, respectively (Figure 18).

8.3.2. [^{3}H] Analysis. The intact U3 RNA, labeled with $NaIO_4$ and [^{3}H]-KBH_4 methods, was digested with RNase A and

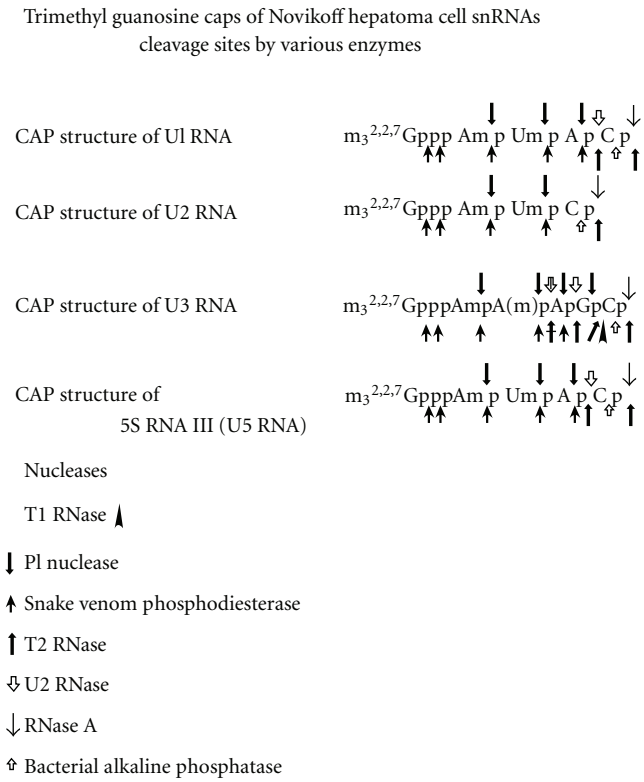


Figure 17: Characterization of $m_3^{2,2,7}G$ caps of snRNAs U1, U2, U3, and U5. The enzyme susceptible bonds are indicated with arrows. The split arrows indicate that some bonds without 2′-O-methylation can be cleaved but the ones with 2′-O-methylated ribose are not.

Table 11: The analysis of [^{32}P]-labeled U2 RNA 5′ oligonucleotide [12]. The 5′ ends obtained from uniformly [^{32}P]-labeled U2 RNA were digested with T1 RNase or RNase A and isolated by two-dimensional electrophoresis (cellulose acetate at pH 3.5 followed by DEAE paper electrophoresis). The 5′ oligonucleotides were digested with snake venom phosphodiesterase before and after removal of 3′ phosphate with bacterial alkaline phosphatase. The products were separated as in Figure 27. The radioactivity ratios are listed.

5′ fragment from U2 RNA	^{32}P ratio in nucleotide							
	Pi	pUm	pAm	$pm_3^{2,2,7}G$	pC	pCp	pG	pGp
5′ oligo from T_1 digestion	1.58	1.33	0.91	0.90	1			1.48
Alkaline phosphatase digested $T_1$5′ oligo	1.22	1.43	0.98	0.90	1		0.93	
5′ oligo from RNase A digestion	1.74	1.25	1	1.08		1.50		
Alkaline phosphatase digested A 5′ oligo	0.93	1.45	0.94	0.77	1			

chromatographed on DEAE-Sephadex (Figure 15). Subsequent digestion by T_1 RNase released only one nucleotide from the RNase A oligonucleotide, indicating that the G was adjacent to a RNase A susceptible pyrimidine. The purified 5′ oligonucleotide obtained after T_1 RNase and RNase A was digested with snake venom phosphodiesterase followed by alkaline phosphatase. The nucleosides and trialcohol derivatives were separated by TLC (Figure 30). The trialcohol derivative of $m_3^{2,2,7}G$ indicates that this nucleotide has free 2′ and 3′ OH at the end of the intact molecule.

8.3.3. [^{32}P] Analysis. The [^{32}P]-labeled U3 RNA digested by T_1 RNase and U_2 RNase was separated by two-dimensional electrophoresis (Figure 29). The enzyme-resistant oligonucleotides 11A and 11B were eluted from the paper and treated with alkaline phosphatase. The products were chromatographed on DEAE-Sephadex A-25 with GMP, GDP, and GTP markers. The 11A (cap II) was eluted at GTP region and 11B (cap I) was eluted at GDP region, indicating that 11B has one nucleotide less than 11A (Figure 31).

From these data, obtained by UV, [^{3}H], and [^{32}P] experiments, the U3 RNA 5′ oligonucleotide sequence has been deduced to be **$m_3^{2,2,7}$GpppAmpA(m)pApGpCp**.

8.4. 5S RNA III (U5 RNA) 5′ End Oligonucleotide. The oligonucleotide sequence was deduced as in the case of U1 RNA. The structure is identical to the U1 5′ oligonucleotide **$m_3^{2,2,7}$GpppAmpUmpApCp**.

Table 12: Tgs1 interacting proteins [70–72]. The Tgs1 (trimethylguanosine synthase 1) interacting proteins are listed. It is interacting with proteins involved in many aspects of RNA metabolism such as transcription, spliceosome assembly, maturation, and modification.

Tgs interacting proteins [70]	
Transcription apparatus (regulators of RNA polymerase II transcription) (1) Rpn4 (TF; Proteasome subunits and U2 snRNA gene) (2) Spt3 (SAGA complex) (3) Srb2 (mediator complex) (4) Soh1 (Med31; mediator complex) (5) Swr1 (ATPase; binds to 5′ end of yeast transcription units) (6) Htz1(H2AZ) (binds to 5′ end of yeast transcription units) (7) CBP (binds to human PIMT/Tgs1) (8) P300 (binds to PIMT/Tgs1) (9) PBP (binds to PIMT/Tgs1) (10) PRIP (binds to PIMT/Tgs1; [71])	Spliceosome assembly (1) Mud1 (yeast homolog of U1A) (2) Mud2 (yeast homolog of U2AF65) (3) Nam8 (Mud15) (homolog of TIA-1) (4) Brr1 (snRNP) (5) Lea1 (U2A; U2snRNP) (6) Ist3 (Snu17; U2 snRNP) (7) Isy1 (Ntc30; interacts with Prp19) (8) Cwc21 (component of Cef1 complex) (9) Bud13 (Cwc26) (Cef1 complex) (10) SMN (HeLa cell) [72]
RNA end processing and decay (1) Trf4 (Pap2; catalytic subunit of TRAMP) (2) Lsm1 (decapping complex) (3) Pat1 (decapping complex)	RNA modifying factors (1) SmB (snRNP; interacts with YNL187/Swt21) (2) SmDL (snRNP) (3) Cbf5 (snoRNP) (4) Nop58 (snoRNP) (5) Mrm (RNA 2′-O-methyltransferase)

9. RNA Signature Modifications for Different RNA Classes

9.1. End Modifications

9.1.1. 5′ End

(a) According to Chemical Nature of Caps

5′ Trimethylguanosine cap for the snRNA,

5′ 7-mehtylguanosine cap for the mRNA,

5′ 2,7 dimethylguanosine cap of virus and nematode RNAs

5′ mpppG of U6 RNA.

(b) According to Flanking Nucleotide Modification of Caps. (See Table 13).

(c) 5′ End Uncapped RNA. (pppNp) for primary transcripts such as 4.5S RNA I, 5S RNA, and Alu RNA. (pNp) 5′ end for processed RNAs such as Alu RNA, 5S RNA, tRNA, YRNA.

9.1.2. 3′ End

3′2′-O-methylated; 4.5S RNA III

3′ poly-A; mRNA, lncRNA

3′ poly-U; polymerase III transcripts such as 4.5S RNA I, 5S RNA, and others

3′ CCA; tRNA, U2 RNA.

9.2. Internal Modifications. The most colorful modifications are in tRNAs that contain methyl, formyl, acetyl, isopentyl, threonyl, carbamoyl, and other groups and modifications by pseudouridylation, deamination, reduction, or thiolation. Focusing on recent findings for snRNA, $m_3^{2,2,7}G$ capping reactions are very interesting because trimethylguanosine is found only in noncoding RNA cap structures, although some nematode mRNA species also contain $m_3^{2,2,7}G$ caps. The snRNAs are less abundant (10^5 copies) than ribosomal RNA or tRNA (10^6 copies). Isolating large amount of RNA can be a hurdle to overcome. Massive preparative procedures and syntheses were pivotal for the thorough analysis of these modifications. The 2′-O-modifications occur mostly internally, and 3′ Um was also found in 4.5S RNA III. The RNA ribose with 2′-O-methylation confers resistance to enzymatic digestion by such enzymes as RNase A, RNase T_1, RNase U_2, and RNase T_2. They are also resistant to alkaline hydrolysis, and the alkaline hydrolysates can be separated into di-, tri-, and tetranucleotides by column chromatography and then by two-dimensional paper chromatography (Figure 16). Other enzymes which can cleave 2′-O-methylated nucleotides are snake venom phosphodiesterase, P1 nuclease, and spleen phosphodiesterase. These are valuable tools for sequencing.

10. Presence of $m_3^{2,2,7}G$ Caps in RNA Species

10.1. Nucleolar RNA. Initially, the $m_3^{2,2,7}G$ cap containing snoRNA was found in U3 RNA [36]. Since then C/D snoRNA and H/ACA snoRNA have been discovered exponentially. The snoRNAs are transcribed from monocistronic as well

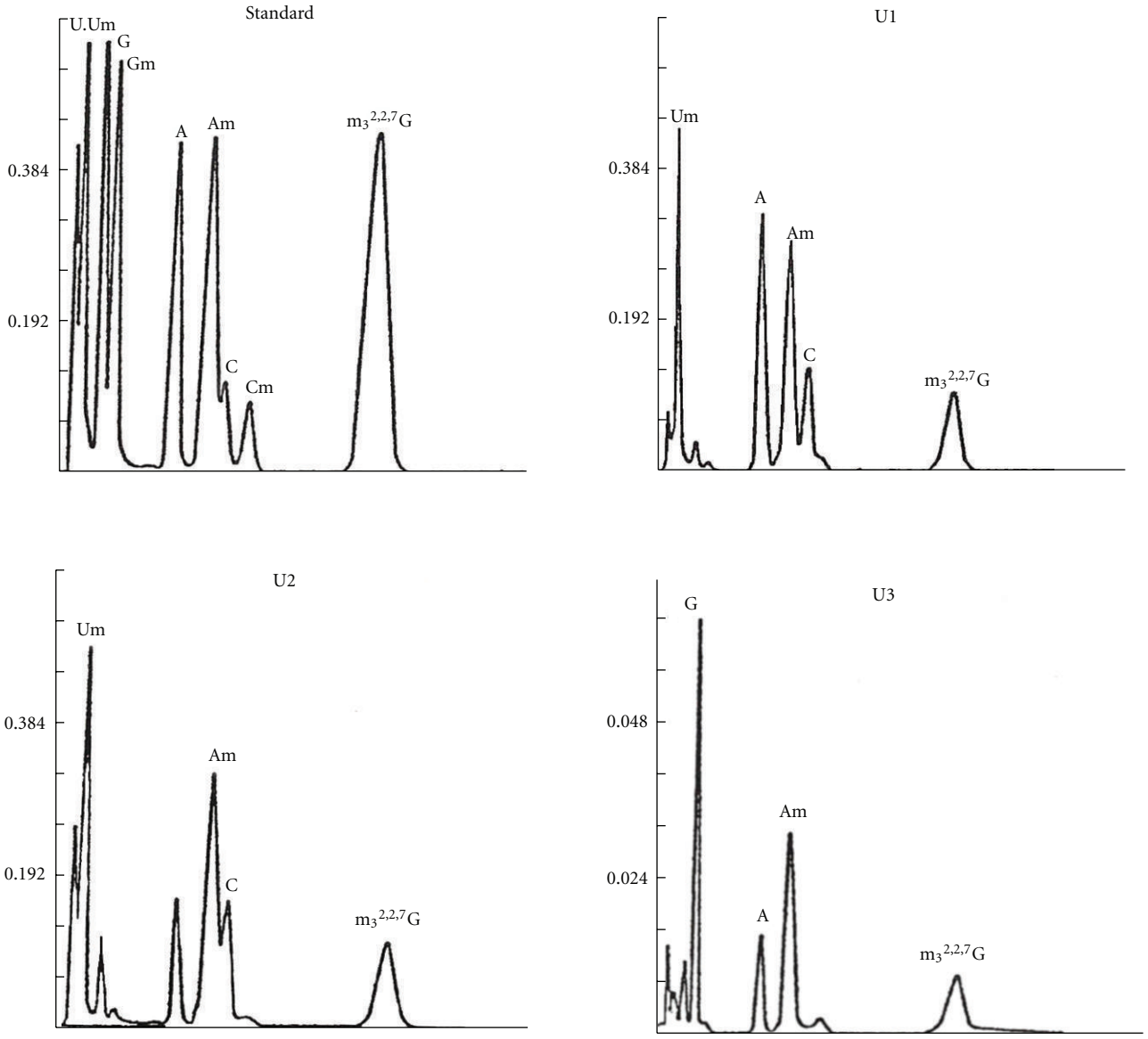


Figure 18: HPLC (High Pressure Liquid Chromatography) pattern of nucleosides of 5′ oligonucleotides of U1 RNA, U2 RNA, and U3 RNA [37]. The 5′ oligonucleotides of U1, U2, and U3 RNAs were obtained by digestion of T_1 RNase and RNase A. The fragments were digested with snake venom phosphodiesterase and alkaline phosphatase. The nucleosides produced were separated by high pressure liquid chromatography (Varian Aerograph Liquid Chromatograph LCS-1000) at 55°C, 700–800 p.s.i. with 0.4 M ammonium formate (pH 3.5). Absorbance at 254 nm was recorded.

as polycistronic independent positions as well as intronic regions of mRNA, especially the genes coding ribosomal proteins. In vertebrates, there have been >76 snoRNAs that have been reported, but only U3, U8, and U13 snoRNAs have been reported to have $m_3^{2,2,7}$G caps [33, 88]. In yeast, there are at least 17 $m_3^{2,2,7}$G cap containing snoRNAs out of more than 76 snoRNAs. It was also reported that some snoRNA precursors, such as pre-snoRNAs 50, 64, and 69, have the $m_3^{2,2,7}$G cap, but mature snoRNA 50, 64, and 69 do not have $m_3^{2,2,7}$G caps. The maturation process cleaves the 5′ fragment by Rnt1 (RNase III like enzyme), and trimming is performed by 5′ → 3′ exonuclease Xrn1 and Rat 1 [89].

10.2. Spliceosomal snRNAs. These include U1, U2, U4, U5, and U6 snRNAs. All of these except U6 contain the $m_3^{2,2,7}$G cap, and U6 has the mpppG cap instead. They are present in complexes as RNP with proteins specific for each RNA as well as some common snRNP proteins such as the Sm proteins. Functionally, U1 RNP acts at 5′ splice sites and U2 RNA at branch sites including 3′ splice sites. U4, U5, and U6 snRNAs enter the spliceosomal intermediate as a tri-snRNP complex.

10.3. Human Telomerase RNA (hTR). Human telomerase RNA has a structure containing the H/ACA motif with 8 conserved regions (CR 1–8) [92].The CR7 contains the CAB

Table 13: Cap variations of flanking nucleotide modifications.

	Trimethylguanosine cap	7-Methylguanosine cap
Type 0	$m_3^{2,2,7}$GpppN-	m^7GpppN-
Type I	$m_3^{2,2,7}$GpppNmN-	m^7GpppNmN-
Type II	$m_3^{2,2,7}$GpppNmNmN-	m^7GpppNmNmN-
Type III	$m_3^{2,2,7}$GpppNmNmNmN-	m^7GpppNmNmNmN-
Type IV	$m_3^{2,2,7}$GpppNmNmNmNmN-	m^7GpppNmNmNmNmN-

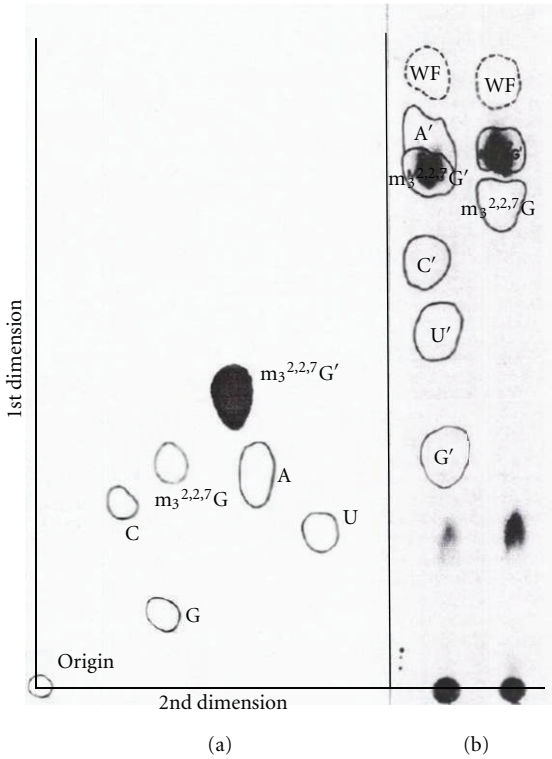


Figure 19: Identification of $m_3^{2,2,7}$G′ [16]. (a) Fluorogram of two-dimensional thin layer chromatography of [^{3}H]-labeled nucleoside trialcohol derivative (N′) released from U1 RNA 5′ fragment. It was identified as $m_3^{2,2,7}$G′ with standard. (b) Fluorogram of chromatographic separation of [^{3}H]-labeled nucleoside trialcohol derivative on Whatman 3MM paper. The [^{3}H]-labeled compound was identified as $m_3^{2,2,7}$G′ with standard.

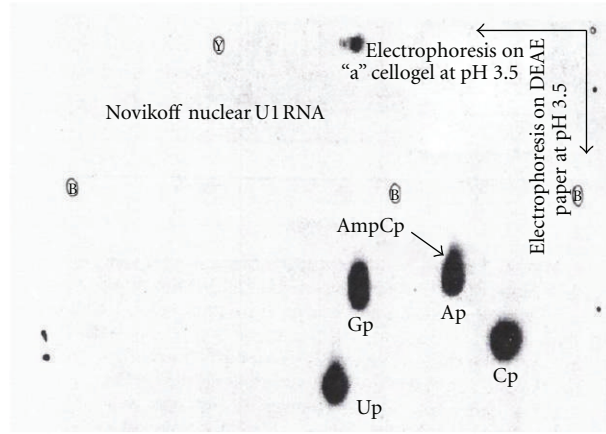


Figure 20: 2D map of U1 RNA digested with T_2 RNase and U_2 RNase [16]. The U1 RNA uniformly labeled with [^{32}P] was digested with T_2 RNase and U_2 RNase. The resistant 5′ fragment (spot "a") was separated from the rest of the hydrolysate by two-dimensional electrophoresis. The first dimension was on cellogel at pH 3.5, and the second dimension was on DEAE paper in 5% acetic acid-NH_4 acetate at pH 3.5.

box (Cajal body box) consensus sequence of UGAG and directs the RNA localization into the CB (Cajal body). The Tgs1 (trimethyl guanosine synthase) is also present in the Cajal body and may be responsible for the $m_3^{2,2,7}$G cap formation. Not all Cajal bodies contain the hTR, and it may be a transient localization for the maturation of hTR in the Cajal body. In the absence of Tgs1, the telomere of yeast *S. cerevisiae* has elongated single-stranded 3′ overhangs and TLC1 (1200 nt telomerase RNA) lacks the $m_3^{2,2,7}$G cap. The absence of Tgs1 causes premature aging of yeast [93, 94].

10.4. C. elegans SL RNA. C. elegans has mRNA with the m^7G cap as well as $m_3^{2,2,7}$G cap, and the expression is regulated differentially. The genes for protein coding are monocistronic as well as polycistronic, and introns are much smaller than observed in mammalian cells. The polycistronic genes contain 2–8 operonic genes regulated by the same promoters. Some gene products are not processed, and others are spliced by cis-splicing as well as transsplicing. The transsplicings are carried out by SL RNA 1 or SL RNA 2. The approximately 110 SL RNA 1 genes are in tandem in chromosome V. The SL RNA 2 is derived from SL RNA 1 and there are ~18 dispersed genes with a variety of variant SL2 RNAs (some are called SL3, SL4, etc.). They are all 100–110 nucleotide long and contain $m_3^{2,2,7}$G caps and Sm protein binding sites. These pre-mRNAs, containing 5′ outron (monocistronic and 5′ first gene in polycistronic operonic genes), are transspliced by SL RNA 1 and internal operonic pre-mRNAs are mostly transspliced by SL RNA 2 and these genes have typically U-rich sequence containing ~100 bp spacers between two cleavage sites. The internal

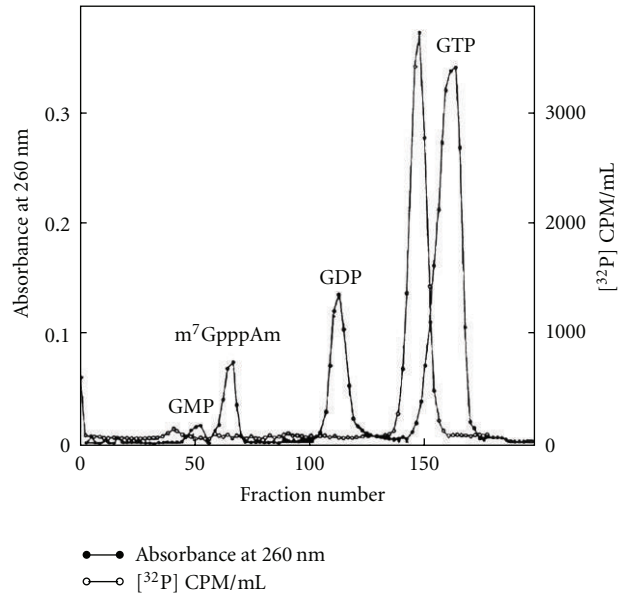


Figure 21: DEAE column chromatography of U1 5′ oligonucleotide [16]. Spot "a" in Figure 20 was eluted and digested with alkaline phosphatase. The product was chromatographed on DEAE-Sephadex A-25 with GMP, GDP, and GTP. The fragment chromatographed at GTP region.

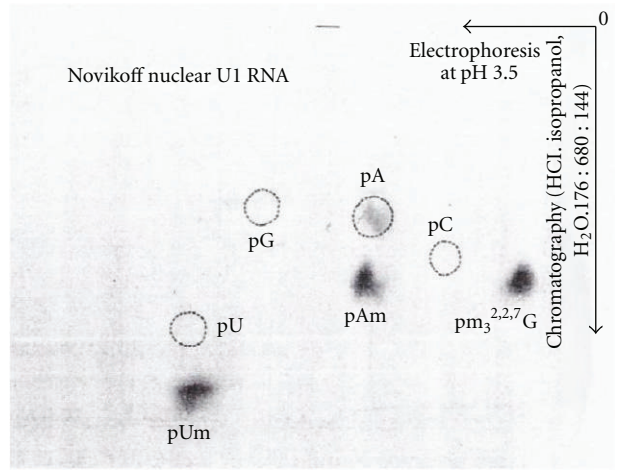


Figure 22: Nucleotide composition of U1 5′ oligonucleotide [16]. The peak from the GTP region was digested with snake venom phosphodiesterase and separated on Whatman 3MM paper by electrophoresis at pH 3.5 (5% acetic acid adjusted pH to 3.5 with ammonium hydroxide) and chromatography in the second dimension with a solvent system consisting of isopropyl alcohol, HCl, and H_2O in the ratio of 680 : 176 : 144 by volume. Autoradiography was performed using X-ray film.

mRNA gene of polycistronic operonic genes, lacking a spacer, is transspliced always by SL RNA I [95, 96]. The transspliced mRNA contains a $m_3^{2,2,7}$G cap containing 22 nucleotides of SL RNA at their 5′ ends. The SL RNA (splice leader RNA) has a $m_3^{2,2,7}$G cap and Sm protein binding sites. The nematode *C. elegans* has 5 eIF4E isoforms of cap binding proteins. They are IFE-1 (m^7G cap and $m_3^{2,2,7}$G cap binding), IFE-2 (m^7G cap binding, but competed by the $m_3^{2,2,7}$G cap), IFE-3 (m^7G cap binding only), IFE-4 (m^7G cap binding only), and IFE-5 (m^7G cap and $m_3^{2,2,7}$G cap binding). The homolog amino acids W56 and W102 stacking the m^7G caps in mice eIF4E are W51 and W97 in IFE-3 and W28 and W74 in IFE-5 (Figure 32).

The differences in 3-4 loop configuration between IFE-5 and IFE-3 are N64Y/V65L. The changes in IFE-5 amino acid asparagine 64 to tyrosine and valine 65 to leucine change binding properties more to m^7G cap binding than to $m_3^{2,2,7}$G cap binding. IFE-5 has 4 cysteines, and its conformation is governed by disulfide bond formation. It is suggested that the cap binding cavity is altered to produce a smaller cavity that discriminates against the $m_3^{2,2,7}$G cap binding [85]. These may provide translational regulation of m7G cap mRNA and transspliced $m_3^{2,2,7}$G cap mRNA in *C. elegans*.

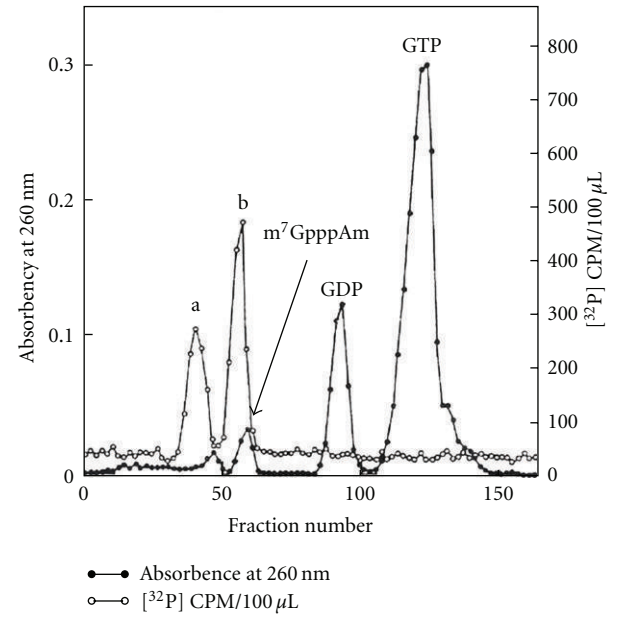
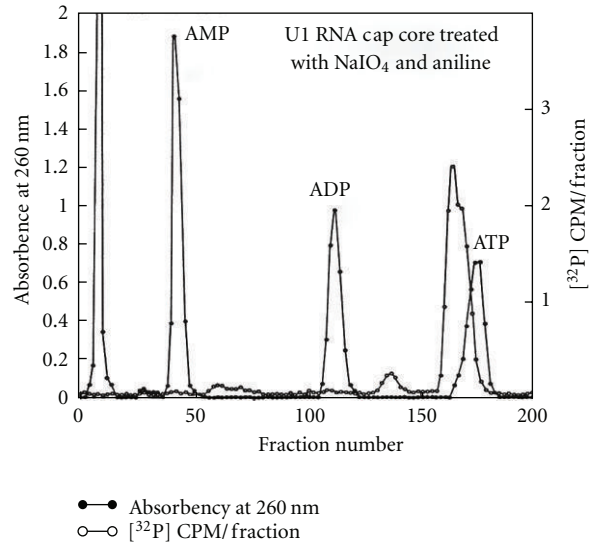


Figure 23: Cap core chromatography [16]. The 5′ oligonucleotide eluted from the GTP region (Figure 21) was digested with P1 RNase and chromatographed on a DEAE-Sephadex A-25 column. Two peaks "a" (mononucleotides pUm and pA) and "b" (cap core $m_3^{2,2,7}$GpppAm) were observed.

11. Synthesis of $m_3^{2,2,7}$G Cap

modifications. Trimethyl G caps are present in snRNAs involved in splicing and also in snoRNA involved in rRNA processing and modifications such as Ψ formation (H/ACA snoRNA) or 2′-O-methylation (C/D snoRNA). These include U1, U2, U4, and U5 spliceosomal RNAs, and U3, U8, and U13 nucleolar RNAs. Recently, telomerase RNA (*S. cerevisiae* TLC1) has also been reported to have



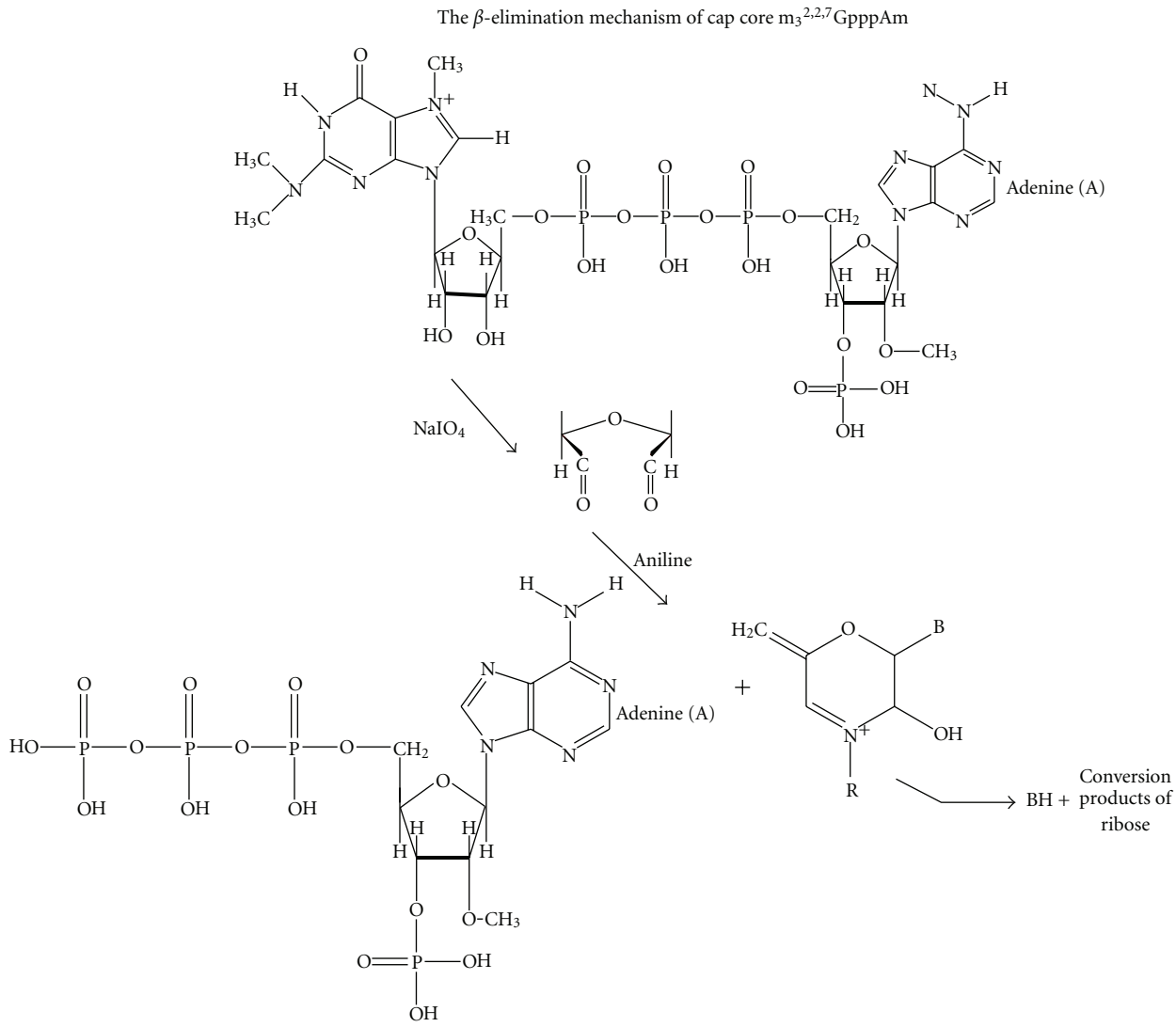


FIGURE 24: The β-elimination of cap core [16]. The cap core (peak b from Figure 23) was treated with $NaIO_4$ and aniline to remove $m_3{}^{2,2,7}G$ by β-elimination reaction. The remaining nucleotide was chromatographed in the ATP region indicating it is pppAm. This proved that the cap core was $m_3{}^{2,2,7}GpppAm$.

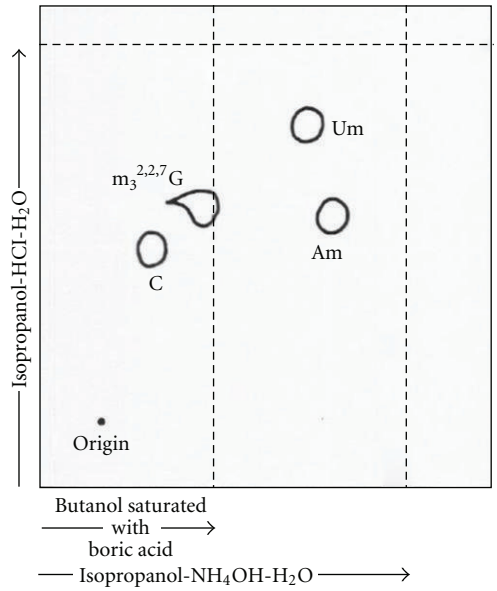


Figure 25: Two-dimensional chromatography in borate system [12]. The nucleoside mixture from U2 RNA 5′ oligonucleotide (RNase A product) was obtained by digestion with snake venom phosphodiesterase and alkaline phosphatase. The resulting nucleosides were separated with the borate system. Um and Am migrated through the butanol-boric acid while the $m_3^{2,2,7}G$ and C, which form borate complexes, were retained in the butanol-boric acid phase.

a trimethylguanosine cap structure. The trimethyl-G caps are formed on cap 0 or cap I of m^7G caps of pre-snRNAs by dimethylation of N2 position by trimethylguanosine synthase (Tgs1). The Tgs1 has been found to be in the Cajal body and cytoplasm. The U3 snoRNA is hypermethylated in the Cajal body, and U1, U2, U4, and U5 snRNA have been reported to be hypermethylated in the cytoplasm.

11.1. The m^7G Cap Formation. The RNA polymerase initiates the RNA transcription with 5′ triphosphate nucleotides and in a majority with purine nucleotides of ATP or GTP. The capping reaction in a polymerase II system occurs cotranscriptionally within the nascent transcript of ~30–50 nucleotides. The guanylyltransferase is attached to heptad (YSPTSPS) repeats of CTD of RNA polymerase II. It was reported with cloned mouse guanylyltransferase and synthetic heptad repeats that the serine 5 phosphorylated 6 heptad repeats stimulated guanylyltransferase activity 4-fold. Serine 2 phosphorylation also binds the guanylyltransferase but did not stimulate enzyme activity [97]. The capping enzymes contain RNAtriphosphatase and RNA guanylyltransferase in the same molecule, but methylating enzymes are in different protein and occurs in separate steps.

The enzymes involved are RNA triphophosphatase and RNA guanylyltransferase, which can be found in the same enzyme, catalyze removal of one phosphate from pppNp initiation nucleotide, and transfer GMP from GTP through intermediary GMP-lysine phosphamide enzyme complex. The RNA guanyl 7 methyltransferase methylates the guanine at N7 position. The RNA 2′-O-methyltransferase methylates penultimate nucleotide 2′ OH, producing the cap 1 structure. In rat liver, it has been reported that 2′-O-methylation may precede the guanosine N7 methylation [98].

The capping reactions by mammalian and shrimp capping complexes (HeLa cell, rat liver, calf thymus, and shrimp) [98] have been reported as below:

RNA Triphosphatase and Guanylyltransferase. The monomer of the 69–73 kDa protein has functions of RNA triphosphatase and RNA guanylyltransferase activity.

$$\text{pppNpNpNpNp-} \xrightarrow{\text{RNA triphosphatase}} \text{ppNpNpNpNp-} \quad (1)$$

$$\text{GTP} + \text{RNA guanylyltransferase} \xrightarrow{Mg^{++}} \text{GMP-(phosphamide)-E} + \text{ppi} \quad (2)$$

$$\text{GMP-E} + \text{ppNpNpNpNp-} \xrightarrow{Mg^{++}} \text{GpppNpNpNpNp-} + \text{RNA guanylyltransferase} \quad (3)$$

$$\text{GpppNpNpNpNp-} + \text{AdoMet} \xrightarrow{\text{RNA 2′-O-methyltransferase}} \text{GpppNmpNpNpNp-} + \text{AdoHcy} \quad (4)$$

$$\text{GpppNmpNpNpNp-} + \text{AdoMet} \xrightarrow{\text{RNA guanyl 7-methyltransferase}} m^7\text{GpppNmpNpNpNp} + \text{AdoHcy} \quad (5)$$

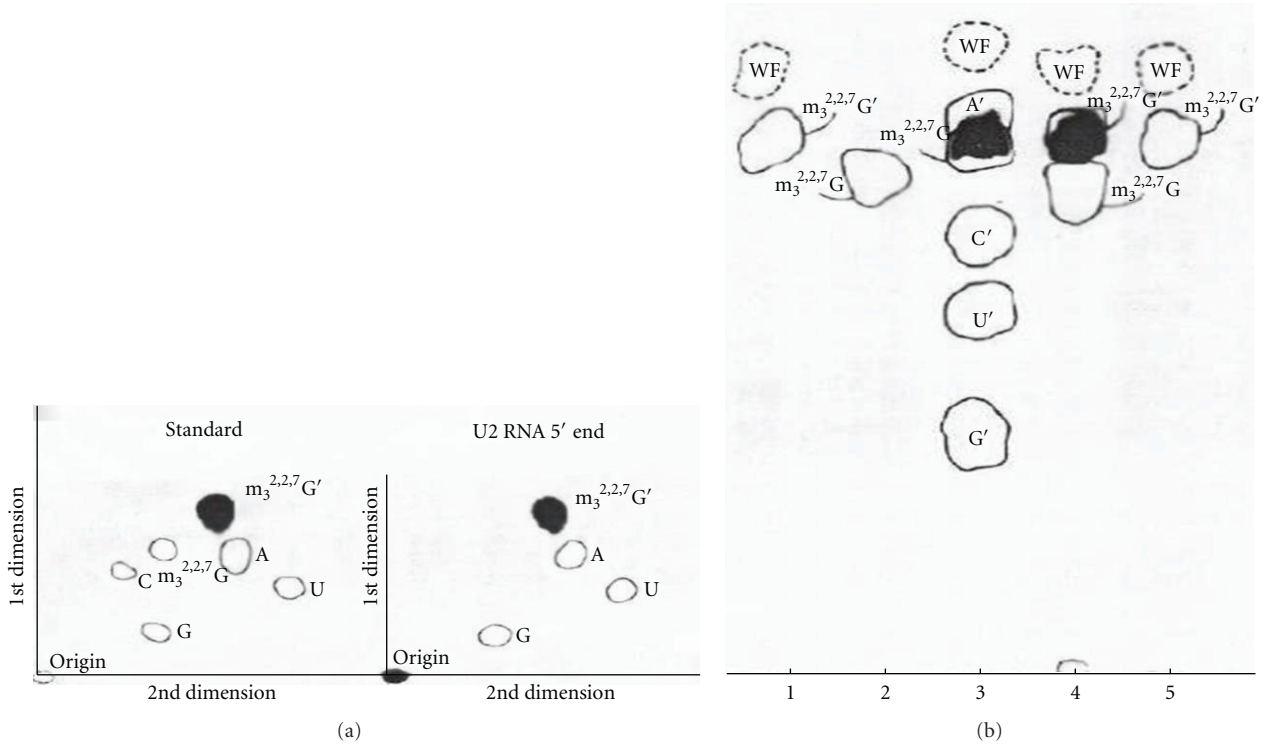


FIGURE 26: Fluorograph of [^{3}H]-labeled trialcohol derivatives of $m_3^{2,2,7}G$ from U2 RNA 5′ end [12]. (a) Two-dimensional thin layer chromatography. (b) One-dimensional paper chromatography. In both systems, the labeled compound was $m_3^{2,2,7}G'$. (See text).

Some of the capping enzymes (vesicular stomatitis virus, spring viremia of carp virus) use the substrate monophosphorylated 5′ end (pNpNpNpNp-) [99, 100], and 7-methylation occurs after the 2′-O-methylation has taken place.

Trimethylguanosine cap synthesis is carried out by multiple steps involving From HeLa cells, two enzymes forming cap I from cap 0 and cap II from cap I have been purified and characterized [101].

11.2. Cap I Methyltransferase. This enzyme is present in both the nucleus (29.3 units/mg) and cytoplasm (3.74 units/mg) and cap II methyltransferase is exclusively in the cytoplasm (4.62 units/mg). Cap I methyltransferase uses $GpppA(pA)_n$, $m^7GpppA(pA)_n$, $m^7GpppApGp$, $m^7GpppApGpUp$, and RNA with type 0 cap as substrates but not m^7GpppA or GpppA. The substrate required for cap I formation should be at least a trinucleotide.

The order of 7-methylation of ultimate G nucleotide and 2′-O-methylation of penultimate nucleotide is uncertain, and both pathways may occur.

11.3. Cap II Methyltransferase. This enzyme is present only in the cytoplasm and converts cap I to cap II. The mature mRNA with 5′ m^7G cap and 3′ polyadenylation is then transported into the cytoplasm as a complex with CBC20/80, PHAX, and Crm1-RanGTP. The m^7G cap binds to CBC20 (156 amino acids) in complex with CBC80 (790 amino acids). The crystal structure of the CBC20/80 complex in association with m7G cap has been reported [86, 87]. The CBC20 is in an unfolded form in the absence of CBC80. The CBC80 has 3 domains, each containing consecutive 5-6 helical hairpins resembling the MIF4G (middle domain of eIF4G). The CBC20 has a typical RRM motif and binds between domains 2 and 3 of CBC80. The m^7G cap is sandwiched between Tyr 43 and Tyr 20. And Phe 83, Phe 85, and Asp 116 have essential role for m^7G cap binding. Asp 116 and Trp 115 interact with the N2 amino group and confer specificity of the m^7G cap for other structures (Figure 33).

In the cytoplasm, the m^7G cap plays a role in the initiation of translation by binding to eIF4E which complexes with eIF4A and eIF4G. The exact mechanism of exchange is not known but CBC80 has binding capacity for PHAX or eIF4G and dissociation of CBC80 from CBC20 makes CBC20 become disordered [86, 87].

11.4. Maturation of snRNAs. The snRNAs synthesized by RNA polymerase II with m^7G cap structures are transported into the cytoplasm in complex with CBP20/80, PHAX (phosphorylated adaptor for RNA export), the CRM1 (export receptor, chromosome region maintenance 1) or exportin 1 and RanGTP (Ras-related nuclear antigen). The snRNPs in the cytoplasm are trimethylated and processed. The mature RNA is reimported into the nucleus in a complex with the trimethyl G cap-specific binding protein snurportin 1 and snRNA binding proteins of Sm RNP and SMN proteins.

Despite immunofluorescent staining of U1 and U2 RNA exclusively in the nucleus [102], biochemical analyses have

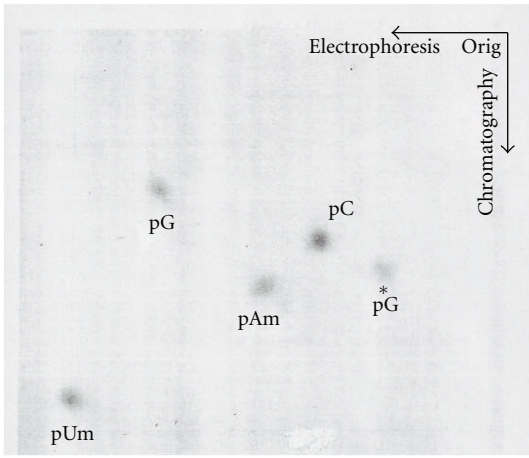


Figure 27: Autoradiograph of nucleotides from [^{32}P]-labeled U2 RNA 5′ fragment [12]. The [^{32}P]-labeled U2 RNA 5′ fragment (T1 RNase digestion) was treated first with alkaline phosphatase and then with snake venom phosphodiesterase. This mixture of mononucleotide products was separated by electrophoresis followed by chromatography. Approximately equal amounts of $pm_3^{2,2,7}G$, pAm, pUm, pC, and pG were observed (Table 11).

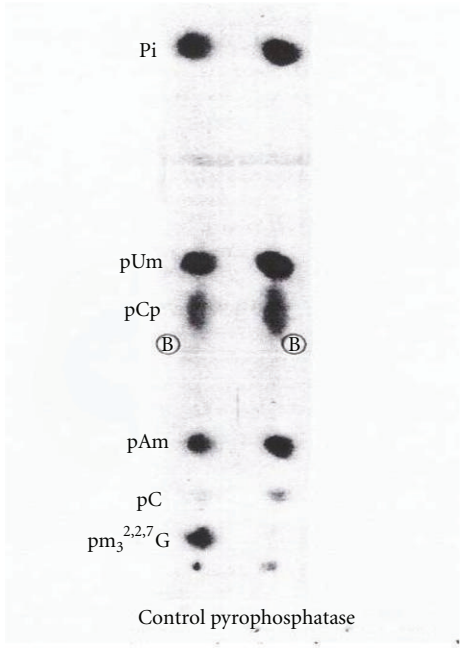


Figure 28: The susceptibility of 5′ cap to pyrophosphatase [12]. The [^{32}P]-labeled 5′ oligonucleotide obtained from U2 RNA by RNase A was digested with pyrophosphatase and base composition was analyzed by snake venom phosphodiesterase digestion. This digestion released $m_3^{2,2,7}G$ from the 5′ fragment indicating that $m_3^{2,2,7}G$ is linked by a pyrophosphate linkage.

demonstrated that trimethylation and maturation of some snRNA takes place in the cytoplasm.

The U1 snRNA [103] and U2 snRNA [104] have been shown to be hypermethylated in the cytoplasm in a Sm protein binding dependent manner. The Xenopus laevis U1 RNA, with the m^7G cap, has been shown to be hypermethy-

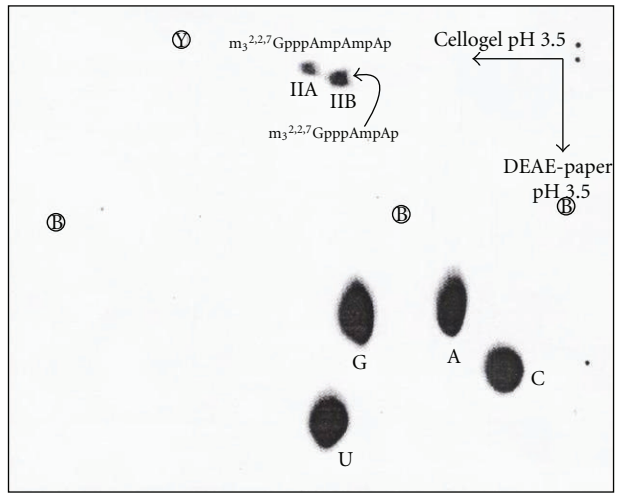


Figure 29: 2D map of U3 RNA digest [58]. The [^{32}P]-labeled U3 RNA was digested with T_2 RNase and U_2 RNase. It produced two 5′ fragments "11A" and "11B".

lated in HeLa cell cytoplasmic extracts and Sm binding site in U1 RNA is required [103]. The Tgs1 has been shown to bind to Sm proteins of Sm B and Sm D. The Xenopus laevis U2 RNA with m^7G cap has been shown to be hypermethylated into the $m_3^{2,2,7}G$ cap structure in enucleated xenopus oocytes [104]. In yeast and human HeLa cells, the Tgs1 for U3 RNA is localized in the nucleolar body of the nucleolus and Cajal bodies, respectively [105]. In the absence of Tgs1 or inactive Tgs1 in yeast, m^7G capped unprocessed U1 RNA is retained in the nucleolus and splicing becomes cold temperature sensitive. The same enzyme is responsible for the U3 nucleolar RNA hypermethylation [106]. The consensus between yeast and human cells is the presence of a nucleolar body in yeast and Cajal body in HeLa cell. The hypermethylation and processing during maturation take place in the nucleolar body in yeast and Cajal body in HeLa cells [105, 106]. The sequence element "UGAG" (also found in the U3 RNA B box) has been reported as a CAB box (Cajal-body-specific localization signal). U3 RNA trimethylation is somewhat different from other snRNAs. The U3 RNA, which does not have Sm protein binding sites, has been shown to require an intact 3′ terminal stem structure for trimethylguanosine cap formation [107].

In HeLa cells, transfected U3 RNA gene products are trimethylated and mature U3 RNA is localized in the nucleolus. Immature U3 RNA, with both m^7G and 3′ extension of 10–15 nucleotides, is detected in Cajal bodies. The nucleolar localization requires the CAB box, hypermethylation to $m_3^{2,2,7}G$ cap, and maturation of the 3′ end [105]. Unlike U1 RNA and U2 RNA, U3 RNA has been shown to be retained in the nuclear compartment and does not go into the cytoplasm for its trimethylation reaction [105, 106, 108].

12. The Tgs1 (Trimethylguanosine Synthase 1)

12.1. Human Tgs1. The Tgs1, trimethylguanosine synthase in human, protein is 110 kDa and 852 amino acids in chain

length. The gene is located in chromosome 8q11. The mRNA is 3.2 kb in length and produces a 110 kDa protein and ~65–70 kDa protein that is proteasome processed. The long form is in the cytoplasm, and the short isoform has been reported to be localized in the Cajal body within the nucleus. The Tgs1 has S-AdoMet methyltransferase signature motifs of X, I, II (include post 1 motif), III, IV, V, and VI [70, 106, 109, 110].

The human Tgs1 motifs are the following.

motif X is (a.a.665)DREGWFSVTPEKIAEHI/FA(a.a.682),

motif I is (a.a.693)VVVDAFCGVGGN(a.a.704),

motif II is (a.a.714)RVIAIDIDPV/IKI(a.a.725) and post 1 motif is VIAID which is responsible for S-AdoMet binding to the enzyme,

motif III is (a.a.740)KIEFICGDFLLLAS(a.a.753),

motif IV is (a.a.758/759)VVFLSPPWGGPDYA(a.a.771/772),

motif V is (a.a.785/786)DGFEIFRLSK(a.a.794/795),

motif VI is (a.a.798/799)NNIVYFLPRNADI(a.a.810/811),

It was reported that trimethylation catalytic activity is located in the C-terminal region (amino acids 631–852) and this region contains the S-AdoMet-dependent methyltransferase motifs. The tryptophan in motif 4 is involved in π stacking with m^7G guanosine of the substrate. The motif 1 and post 1 motif are reported to interact with S-AdoMet. [110]. The C-terminal domain is localized in the Cajal body and binds to C/D-snoRNA- and H/ACA-snoRNA-associated proteins such as fibrillarin, Nop56, as well as dyskerin [110].

The N-terminal portion of the molecule (amino acids 1–~477) has been reported to contain GXXGXXI, a K-homology domain for RNA binding, and a motif for SmB and SmD1 binding. The Tgs1 has also been shown to interact with PRIP (proliferator-activated receptor-interacting protein), and the N-terminal portion (amino acids 1–384) of Tgs1 has been shown to have stimulatory effects on transcription of PPARγ and RXRα [109, 110].

The human Tgs1 (618–853) has been crystallized for structural analysis. The one monomer consists of 11 α-helices and 7 β-strands. It is composed of 2 domains, the core domain (Glu675-Asp844) and N-terminal extension (Leu34-Ser671) connected by 3 amino acids—Val672, Thr673, and Ser674. The core domain consists of 7-β-strands in topology of $\beta6\uparrow\beta7\downarrow\beta5\uparrow\beta4\uparrow\beta1\uparrow\beta2\uparrow\beta3\uparrow$ with a classical class 1 methyltransferase fold resembling the Rossmann-fold AdoMet-dependent methyltransferase superfamily [90]. The N-terminal α-helices form a separate small globular subdomain involved in recognition and binding of both substrates. The residues Glu667 and Phe670 in motif X as well as Pro765, Trp 766, and Pro769 in motif IV are in proximity permitting the top of their binding clefts to be close together. Tryptophan 766 and m^7G are stacked in a coplanar manner with a 3.2 Å distance providing a tight π-π interaction between them (Figure 34).

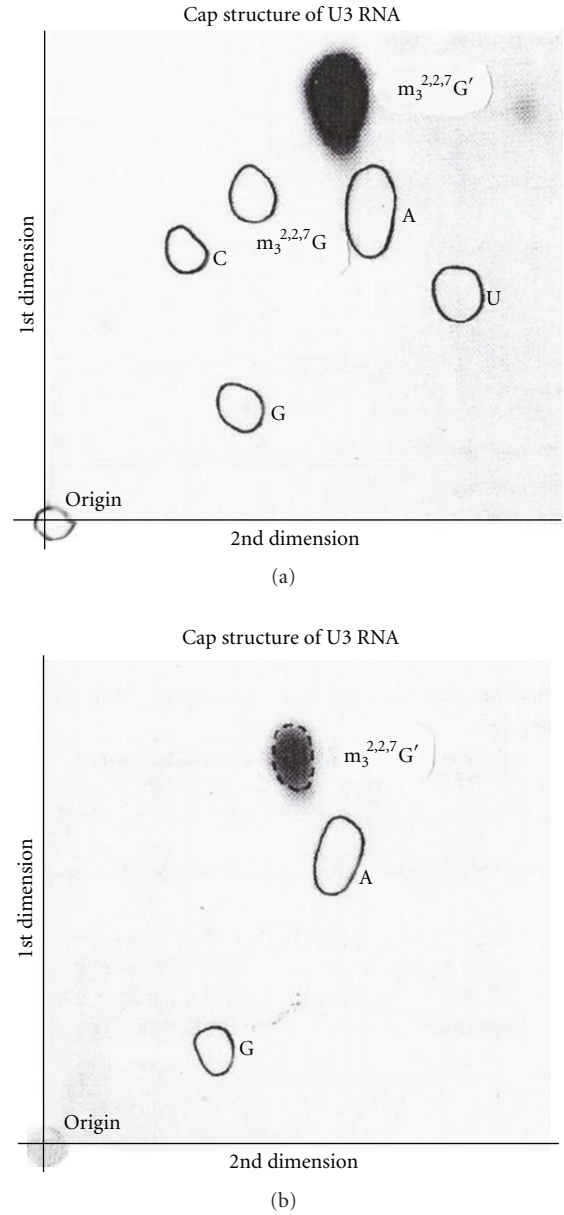


Figure 30: $m_3^{2,2,7}G'$ identification from U3 RNA 5′ end fragment [58]. The [^{3}H]-labeled U3 RNA 5′ fragment obtained by RNase T_1 and RNase A was digested with snake venom phosphodiesterase and alkaline phosphatase. The nucleoside mixture was separated by two-dimensional thin layer chromatography with standard nucleoside mixture in (a) and with the trialcohol derivative of $m_3^{2,2,7}G'$ in (b). The released nucleoside trialcohol derivative was identified as $m_3^{2,2,7}G'$ by fluorography.

The catalytic mechanism of methylation is by an S_n2 substitution reaction. The N2 of m^7G does the nucleophilic attack on an activated methyl group of the AdoMet (Figure 35).

Dimethylation is not processive. After formation of $m_2^{2,7}G$ both products ($m_2^{2,7}G$ and AdoHcy) dissociate from the enzyme. Tgs1 can use $m_2^{2,7}G$ as a substrate, and newly bound AdoMet can methylate at the same position by the same mechanism to form the $m_3^{2,2,7}G$ cap structure.

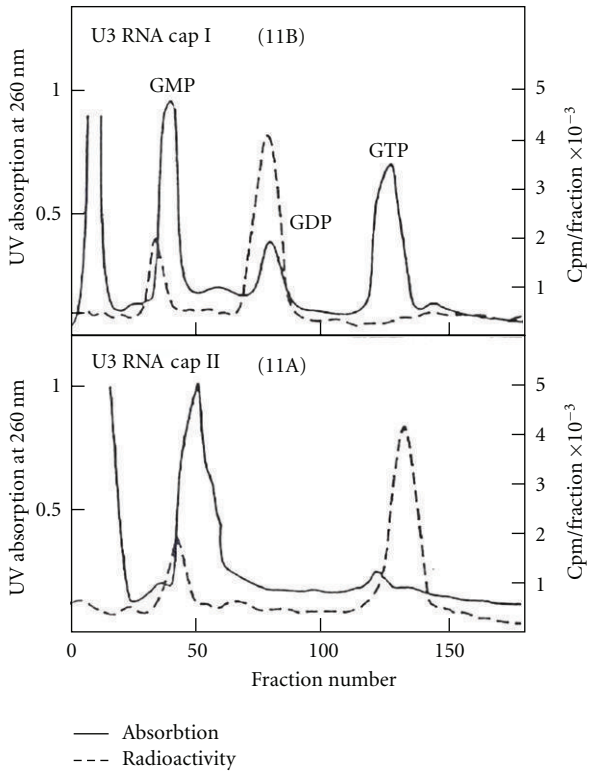


Figure 31: U3 RNA cap I and cap II chromatography [58]. The 5′ fragment from Figure 29 was eluted and treated with alkaline phosphatase and chromatographed on DEAE-Sephadex A-25 with standards of GMP, GDP, and GTP. Component "11B" was eluted in the GDP (cap I, $m_3^{2,2,7}$GpppAmA) region and "11A" (cap II, $m_3^{2,2,7}$GpppAmAmA) was eluted in the GTP region.

12.2. Drosophila Tgs1. In Drosophila melanogaster, DTL (Drosophila TAT-like) has been reported to exhibit trimethylguanosine cap formation activity for both U2 and U4 snRNAs. The mRNA for the protein Tgs1 is polycistronic and 2,600-nucleotide long with upper and downstream ORFs (open reading frames). The uORF is 80 bp from the transcription start site and has coding capacity for a 178 amino acid protein while dORF is 538 bp from the 5′ end and produces a 60 kDa protein (491 amino acids). The two cistrons are overlapped by 76 bp. Mutational analysis indicates that both the uORF and dORF regions are required for viability. The putative product of uORF contains periodic Leu residues, but there is no evidence that this region is translated at any time during Drosophila development. The protein from dORF contains an Arg-rich motif KKKRRQRQI similar to the RNA binding motif RKKRRQRRR in HIV TAT. This protein is localized in the nucleus and responsible for trimethylation of U2 and U4 snRNAs [111].

12.3. Yeast Tgs1. In yeast *S. cerevisiae*, Tgs1 is in the nucleolus and U3 RNA is also in the nucleolus. In the absence of Tgs1, the pre-U3 RNA was found within the nucleolar body and U1 RNA was retained in the nucleolus. *S. cevisiae, S. pombe,* and *Giardia lamblia* Tgs1 can methylate m^7GTP, m^7GDP, and m^7GpppA as substrates without preassembly of snRNP containing Sm proteins. The Tgs1 of S pombe is 239-amino-acid long and m^7G is the pre-requisite for this reaction [112].

12.4. The G. lamblia Tgs1 and Tgs2. The *lamblia* has 2 enzymes, Tgs1 and Tgs2. Tgs 1 is not a processive enzyme but distributive and produces $m_3^{2,2,7}$G in excess of AdoMet and enzyme. However, Tgs2 produces only $m_2^{2.7}$G, and some *G. lamblia* RNAs contain dimethylG caps. The G lamblia Tgs1 has 300 amino acids and Tgs2 is 258 amino acids long. They all have landmark motifs for Ado-Met-dependent methyltransferase motifs [113].

13. Parasite Capping Enzyme *(Trypanosoma brucei)*

The parasite *Trypanosoma brucei* SL RNA (splicing leader) has the biggest 5′ oligonucleotide, type IV, of m^7Gpppm$_2^{6,6}$AmpAmpCmpm^{3}UmpAp [114, 115]. Enzymes involved in the synthesis of this cap structure are TbCgm1, TbCet1, TbMTr1(cap1 2′OMTase), TbMTr2/ TbCom1/TbMT48(cap2 2′OMTase), TbMTr3/ TbMT57(cap3 2′OMTase). However, $m_2^{6,6}$A and m^3Um methylating enzymes have not been identified as yet [115].

13.1. TbCgm1 (T. brucei Cap Guanylyltransferase Methyltransferase 1). There exist two enzyme systems for 5′ cap formation. The first is the system composed of separate independent enzymes which are TbCet1 (*Trypanosoma brucei* triphosphatase, 253 amino acids), TbCe1 (*Trypanosoma brucei* guanylyltransferase, 586 amino acids), and TbCmt1 (*Trypanosoma brucei* m^7G Cap methyltransferase 1, 324 amino acids). The second is a set of fused enzymes possessing dual activities. It is TbCgm1 (*Trypanosoma brucei* cap guanylyltransferase and methyltransferase 1) that has 1050 amino acids [116] with dual activities of guanylyltransferase and guanine N-7 methyltransferase [117]. The TbCe1 guanylyltransferase has 250 amino acids at its N-terminal region which is not found in fungal or metazoan guanylyltransferase and has homology with the phosphate binding loop found in ATP- and GTP-binding proteins [118]. Silencing TbCe1 and TbCmt1 had no effect on parasite growth or SL RNA capping, but TbCgm1 was essential for parasite growth and silencing TbCgm1 increased the amount of uncapped SL RNA. The protein TbCgm1 has guanylyltransferase activity in N-terminal 1–567 amino acids and methyltransferase activity in C-terminal 717–1050 amino acids. The N-terminal guanylyltransferase portion contains 6 colinear guanylyltransferase motifs: I(KADGTR), III(FVVDAELM), IIIa(LIGCFDVFRYVI), IV(DGFIF), V(QLXWKWPSMLSVD), and VI(WSIERLRNDK). The C-terminal methyltransferase por-tion contains regions homologous to m^7G methyltransferase from *T. cruzi* and *L. major* [117].

13.2. Cap Methylating Enzymes: TbMTr1, TbMTr2(TbCom1/ TbMT48), and TbMTr3(TbMT57). They contain a K^{95}-D^{207}-K^{248}-E^{285} tetrad critical for AdoMet-dependent

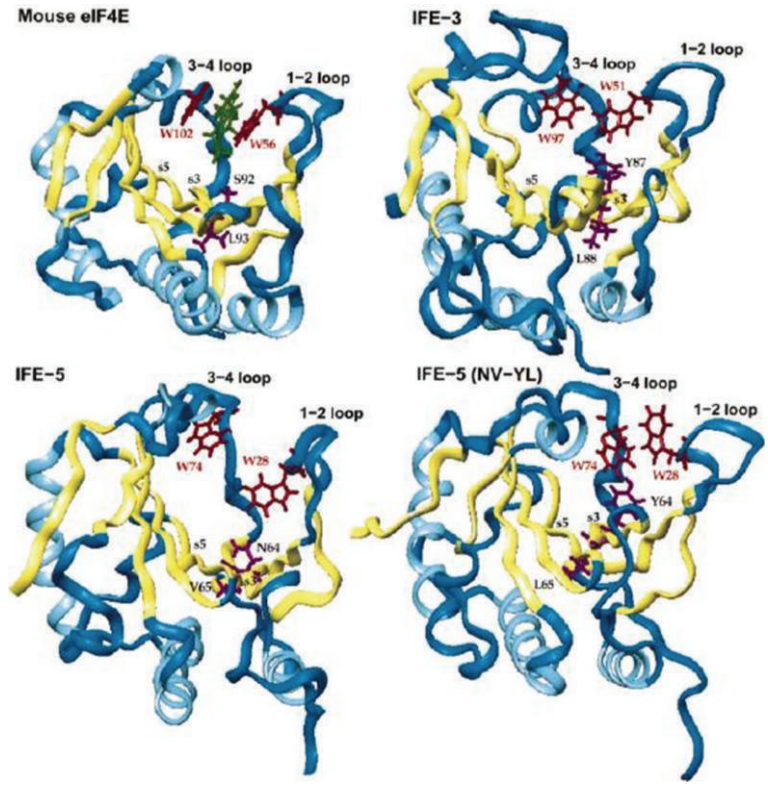


Figure 32: Molecular models of IFE-3, IFE-5, and IFE-5 (NV-YL) in comparison with mouse eIF4E [85]. The differences between the m^7G cap binding pocket and $m_3^{2,2,7}G$ cap binding pocket are illustrated by differences in the 3-4 loop configuration.

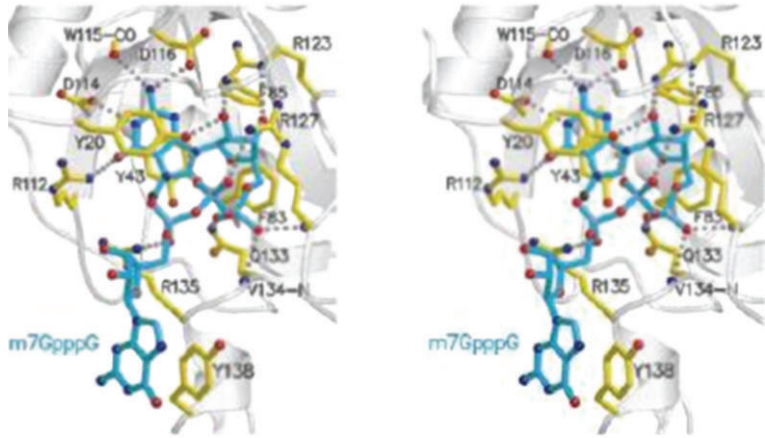
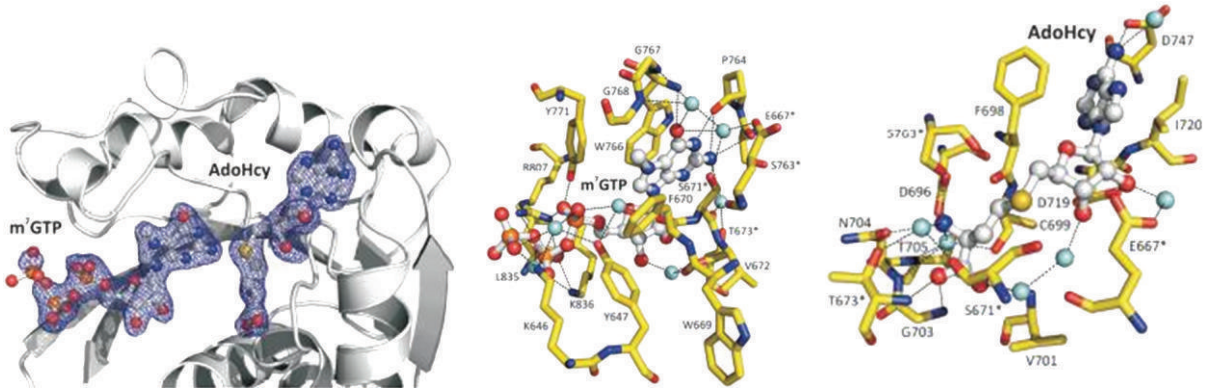


Figure 33: CBP20 (cap binding protein 20) binding to m^7G cap [86, 87]. The m^7G is stabilized by stacking energies between tyrosine 20 and tyrosine 43.

methyltransferase and can convert cap type 0 of *Trypanosoma* SL RNA and U1 snRNA into type 1 cap [115]. The KDKE mediates S_n2 type transfer of methyl groups that involve 2′-OH deprotonation. The U1 snRNA 2′-O-methylation takes place before Sm protein binding to the RNA and it is prerequisite for the dimethylation at the N2 position to make $m_3^{2,2,7}$GpppAm cap structures. Other $m_3^{2,2,7}G$ cap-containing snRNAs such as U2, U-snRNA B (U3 snRNA homolog), and U4 snRNAs were reported to be synthesized by RNA polymerase III in *Trypanosomes*.

The TbMTr2 and TbMTr3 are responsible for second and third nucleotides 2′-O-methylations. The enzymes that perform $m_2^{6,6}A$, m^3U base methylations, and fourth nucleotide 2′-O-methylating enzymes are not known yet.



(a) (b) (c)

FIGURE 34: Crystal structure of hTgs1 [90]. The crystal structure of hTgs1 with substrate m^7GTP and AdoHcy (at the site for AdoMet). The m^7G is stacked between tryptophan 766 and serine 671 [90]. (a) Relative orientation of substrates m^7GTP and AdoHcy at binding pockets. (b) Detailed view of the binding pocket for m^7GTP (shown W766 and S671). (c) Detailed view of the binding pocket for AdoHcy.

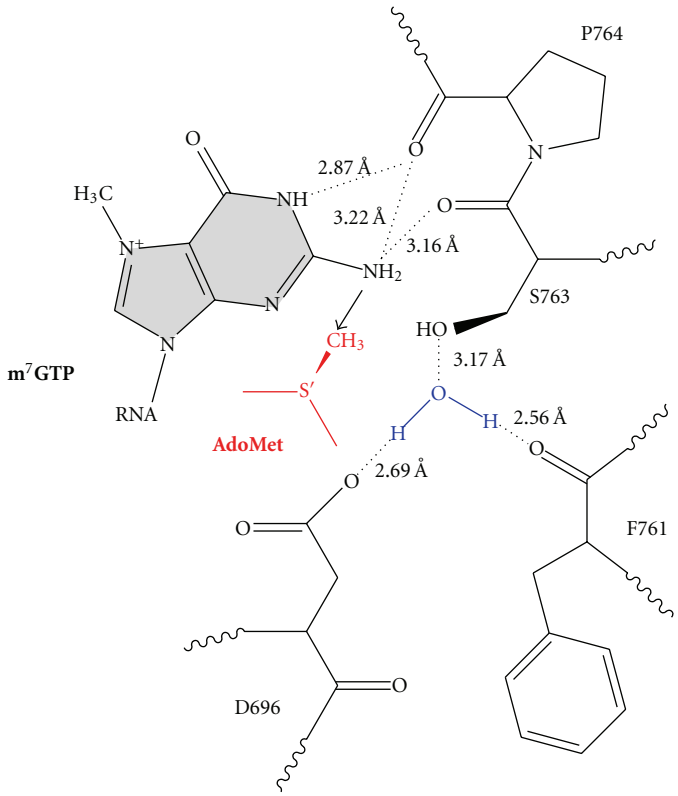


FIGURE 35: The active site of hTgs1 [90]. Proposed mechanism of methyltransferase activity of hTgs1 is shown. The AdoMet methyl group is in close proximity to N2 position of m^7GTP. The prerequisite as a substrate for hTgs1 is m^7G moiety.

14. Transport of Mature RNAs

The snurportin1 is a specific trimethyl G cap binding protein with an importin β binding site at its N-terminus (amino acids 1–65) and trimethyl G cap binding site at amino acids 95–300 forming a cap binding pocket. This protein has more resemblance to mRNA guanylyltransferase. The snurportin 1 binds the trimethyl G cap forming π-stacking with tryptophan 276 and the penultimate purine nucleotide G (Figure 36). The tryptophan 107 is in close proximity to dimethylamine of N2 G suggesting a cation-π interaction and has a role in discriminating between m^7G cap and $m_3{}^{2,2,7}$G cap [91].

15. Tgs 1 Interacting Proteins

Genetic and biochemical analysis of Tgs1 interacting proteins reveals a wide range of proteins involved in RNA metabolism.

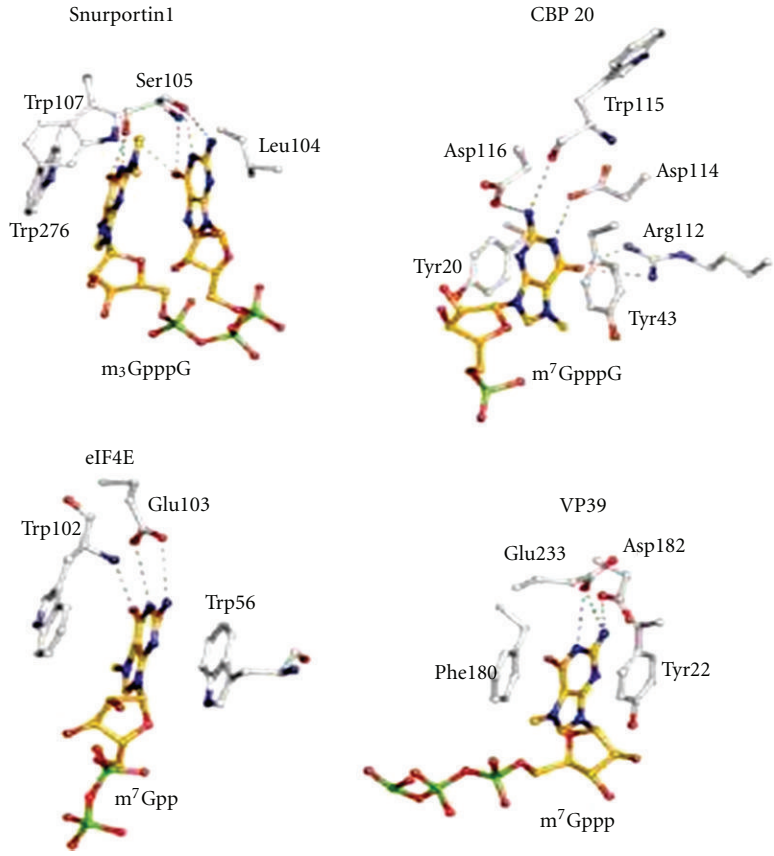


Figure 36: The comparison between $m_3^{2,2,7}G$ binding mode to snurportin1 and m^7G cap binding mode to CBP 20, eIF4E and VP39 [91]. The $m_3^{2,2,7}G$ is stacked between tryptophan 107 and penultimate nucleotide G of cap core $m_3^{2,2,7}GpppG$.

pppGGUCGAGAGG AUGGCUCAGC CGUUAA[**AG**]GC UAGGCCAAAA AUAA{CACCUA
U}AAGAGUUCG GUUCCC[**AG**]CA CCACGGCUGU CCUUCCAGCA CCUUUU-OH

Figure 37: The sequence of 4.5S RNA I (Novikoff hepatoma cell nuclei).

It interacts with proteins in the transcriptional apparatus, RNA end processing and decay, spliceosomal assembly and RNA modifying factors (Table 12).

Structurally, it is distinct from the m^7G cap, and the specificity of binding proteins may determine the precision of its functional role in the RNP complex. The $m_3^{2,2,7}G$ cap structures are present only in nuclear snRNAs and snoRNAs which confer the function within the nucleus in transcription, splicing, modification, processing, and maturation of different RNA species.

16. Conclusion

16.1. General Consideration. In the present postgenomic era, study of the structure and function of noncoding RNAs is supremely important. It is estimated that ncRNAs are probably involved in all aspects of cell metabolism. Therefore, RNA-based information will contribute greatly to understanding various cell metabolisms. In the process of exploring ncRNAs, there may be many surprises awaiting us.

They may include

(1) new species of RNA,

(2) new mechanism of RNA processing,

(3) new mechanism of transcription,

(4) new disease caused by RNAs with pathogenic sequences,

(5) new function for ncRNA.

16.2. The Problem of Unknown Modified Nucleotides. In the process of oligonucleotides cataloging, it is natural that an examination of base composition will reveal modified nucleotides or nucleosides in addition to unmodified standard nucleotides or nucleosides. In routine work, identification of modifications can be readily made by

two-dimensional paper chromatography for nucleotides or thin layer chromatography for nucleosides. However, there may be an occasion where chromatographic identification is not sufficient. Of course, it is best to have collaboration with outside specialists. For the sake of structural microanalysis, it is highly recommendable to determine molecular weight of the unknown nucleotide or nucleoside by mass spectrometry [119]. The required quantity is approximately 5 μg/nt where chromatographic identification of isotopically labeled sample requires 0.5 μg/nt. A difficulty may be confronted with purine bases that are fused to an imidazole ring (Queuosine) which is not suited for mass spectrometry. It is convenient to probe chemical complexity based on mass. The detailed analysis may require an unpredictably large amount of samples. There are 135 modified nucleosides listed, among which 6 nucleosides are not thoroughly identified [1].

16.3. Significance of Sequence Work. Past sequence work has permeated numerous significant areas of research providing a better understanding of cellular metabolism. The information obtained thus far is RNA-based information which is not seen in DNA, proteins, and others. As sequence work continues to make enormous progress, the postgenomic era will shape the direction of research in the area of molecular mechanisms of RNA metabolism. They are briefly as follows.

In RNA maturation, knowledge of structural modifications is necessary to discern between various mechanistic options. For example, there are two molecular mechanisms mediated by catalysis. One is mediated by RNA enzymes (snRNAs and snoRNAs) involved in splicing of pre-mRNA and processing of pre-rRNA. The other is protein enzymes involved in 5′ cap formation. Currently, the higher order structural analysis is in progress. There is a need to elucidate the details of molecular mechanisms.

Along with the study of splicing physiology, splicing pathology is making significant progress. Aberrant modifications can generate disease causing alterations in structure. The aberrations cause problems in reading both genetic codes and splicing codes. Studying the regulation of alternative splicing will clarify the selective rules in intron removal and pathogenic rules in splicing code. From these studies, corrective strategy will evolve. The present sequence work is engaged in definition of ncRNAs diversity and their functional roles [120]. Since it is suggested that ncRNAs are involved in all aspects of regulations in cell metabolism, there may be opportunities to study various paths in cell metabolism, not limited to transcriptional and posttranscriptional events. It is this gigantic task, to reevaluate the genomic work, that holds excitement and promise.

Abbreviations Used in Table 4

Short Noncoding RNA (Usually Shorter Than tRNA and Some Are Longer but Excluding snRNA Such As U1–U13)

miRNA: MicroRNA (imperfect base pairing)

siRNA: Small interfering RNA (perfect base pairing)

tasiRNA: Transacting small interfering RNA

natsiRNA: Natural antisense transcribed small interfering RNA

piRNA: PIWI interacting RNA (RNA precipitated by PIWI protein antibody)

rasi RNA/pitRNA: Repeat-associated small interfering RNA/pi-target RNA

PARs: promoter associated RNAs

PROMTs: Promoter upstream transcripts (sense and antisense transcript)

PASRs: Promoter-associated small RNAs

TSSa-RNAs: Transcription-start-site-associated RNAs

tiRNAs: Transcription initiation RNAs

MSY-RNA: MSY2-associated RNAs (MSY; Y chromosome male-specific protein)

snoRNA: Small nucleolar RNA (C/D box RNA, H/ACA RNA)

sdRNA: sno-derived RNAs

moRNA: MicroRNA-offset RNAs

tel-sRNA: Telomere small RNAs

crasiRNA: Centrosome-associated small interfering RNAs

hsRNA: Heterochromatin small RNA or hairpin small RNA

scaRNAs: Small Cajal-body-associated RNAs

Y RNAs: Cytoplasmic small RNA Y1, Y3, Y4, and Y5

tRNA-derived RNAs: Small RNA processed from tRNA by RNase (angiogenin)

Alu/SINE RNA: Alu restriction enzyme cleaved repeat gene transcript/short interspersed nucleotide element RNA.

Lnc RNA: Long Noncoding RNA (~0.5 to 100 kb)

(1) Specific Long Noncoding RNA

TR/TERC: Telomerase RNA/telomerase RNA component

NEAT RNAs: Nuclear enriched abundant transcript 1 RNAs

NEAT1v-1: NEAT1 variant 1

NEAT1v-2: NEAT1 variant 2

NEAT2/MALAT1: Metastasis associated in lung adenocarcinoma transcript 1

PINC RNA: Pregnancy-induced noncoding RNA

DD3/PCA3: Prostate-cancer-associated RNA 3

PCGEM1: Prostate cancer gene expression marker 1

SPRY4-1T1: Sprouty homolog 4 gene transcript 1 (melanoma specific).

(2) Imprinting-Associated lncRNAs

xiRNAs: X chromosome inactivating RNAs

Xist: X chromosome inactivating sense RNA
Tsix: Antisense transcript of Xist
RepA: Repeat A RNA

AIR RNA: Igf2r imprinting region RNA

H19: Igf2 imprinting region RNA

KCNQ1ot1: Antisense RNA from intron 10 of Kcng1 gene imprinting region.

(3) Regulatory lncRNAs

HOTAIR: Homeogene inactivating RNA

BORG: BMP/OP-responsive-gene-associated RNA

CTN RNA: Cationic amino acid transporter protein coding region RNA

ANRIL RNA: Antisense noncoding RNA in INK4 locus.

(4) Gene-Recombination-Associated lncRNA

LINE: long interspersed nucleotide element

CSR-RNA: Immunoglobulin class switch recombination region RNA.

(5) Satellite DNA Transcripts.

Abbreviations for Table 12

Mud: Mutant U1 die

RES complex: Heterotrimeric RNA retention and splicing complex composed of Bud13, Ist3/Snu17, and Pml1

Swt21: Synthetic with Tgs1 number 21

TRAMP complex (Trf4, Air2, Mtr4p polyadenylation complex): Interacts with exosome in the nucleus and involved in 3′ end processing of rRNA, snoRNA, and U1, U4, and U5 snRNA;

Trf4 or Trf5: poly(A) polymerase(PAP);
Mtr4: RNA helicase;
Air1 or Air2: Zinc knuckle protein

Cbf5(YLR175W): Centromere binding factor;

Pseudouridine synthase catalytic subunit of box H/ACA snoRNP complex

PIMT: PRIP-interacting protein with methyltransferase domain, PIMT is a Tgs1 (trimethylguanosine synthase 1) cloned from human liver cDNA library

PRIP: PPAR interacting protein

PPAR: Peroxisome proliferator-activated receptor

PBP: PPAR binding protein.

Acknowledgment

The authors are indebted to Professor Lynn Yeoman of Baylor College of Medicine for helpful discussion in the preparation of this paper.

References

[1] S. Dunin-Horkawicz, A. Czerwoniec, M. J. Gajda, M. Feder, H. Grosjean, and J. M. Bujnicki, "MODOMICS: a database of RNA modification pathways," *Nucleic Acids Research.*, vol. 34, pp. D145–D149, 2006.

[2] T. A. Cooper, L. Wan, and G. Dreyfuss, "RNA and Disease," *Cell*, vol. 136, no. 4, pp. 777–793, 2009.

[3] J. D. Watson and F. H. C. Crick, "Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid," *Nature*, vol. 171, no. 4356, pp. 737–738, 1953.

[4] I. R. Lehman, M. J. Bessman, E. S. Simms, and A. Kornberg, "Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from Escherichia coli," *Journal of Biological Chemistry*, vol. 233, no. 1, pp. 163–170, 1958.

[5] M. J. Bessman, I. R. Lehman, E. S. Simms, and A. Kornberg, "Enzymatic synthesis of deoxyribonucleic acid. II. General properties of the reaction," *Journal of Biological Chemistry*, vol. 233, no. 1, pp. 171–177, 1958.

[6] M. Muramatsu and H. Busch, "Isolation, composition, and function of nucleoli of tumors and other tissues," in *Methods in Cancer Research*, H. Busch, Ed., vol. 2, Chapter 9, pp. 303–359, Academic Press, New York, NY, USA, 1967.

[7] T. S. Ro-Choi and H. Busch, "Low-molecular-weight nuclear RNAs," in *The Cell Nucleus*, H. Busch, Ed., chapter 5, Academic Press, New York, NY, USA, 1974.

[8] J. Sperling, M. Azubel, and R. Sperling, "Structure and function of the pre-mRNA splicing machine," *Structure*, vol. 16, no. 11, pp. 1605–1615, 2008.

[9] K. Randerath and E. Randerath, "A tritium derivative method for base analysis of ribonucleotides and RNA," *Procedures in Nucleic Acid Research*, vol. 2, pp. 796–828, 1971.

[10] J. A. Gerlt, "Mechanistic principles of enzyme-catalyzed cleavage of phosphodiester bonds," in *Nucleases*, S. M. Linn, R. S. Lloy, and J. R. Roberts, Eds., chapter 1, Cold Spring Harbor Laboratory, New York, NY, USA, 1993.

[11] Y. Motorin, S. Muller, I. Behm-Ansmant, and C. Branlant, "Identification of modified residues in RNAs by reverse transcription-based methods," *Methods in Enzymology*, vol. 425, pp. 21–53, 2007.

[12] T. S. Ro-Choi, Y. C. Choi, D. Henning, J. McCloskey, and H. Busch, "Nucleotide sequence of U2 ribonucleic acid. The sequence of the 5' terminal oligonucleotide," *Journal of Biological Chemistry*, vol. 250, no. 10, pp. 3921–3928, 1975.

[13] L. Cartegni, J. Wang, Z. Zhu, M. Q. Zhang, and A. R. Krainer, "ESEfinder: a web resource to identify exonic splicing enhancers," *Nucleic Acids Research*, vol. 31, no. 13, pp. 3568–3571, 2003.

[14] J. Brachet, *The Biological Role of Ribonucleic Acids*, Weizman Memorial Lecture Series, The Netherlands Elsevier Publishing, 6th edition, 1960.

[15] C. Branlant, A. Krol, J. P. Ebel et al., "Nucleotide sequences of nuclear U1A RNAs from chicken, rat and man," *Nucleic Acids Research*, vol. 8, no. 18, pp. 4143–4154, 1980.

[16] T. S. Ro-Choi and D. Henning, "Sequence of 5' oligonucleotide of U1 RNA from Novikoff hepatoma cells," *Journal of Biological Chemistry*, vol. 252, no. 11, pp. 3814–3820, 1977.

[17] C. Branlant, A. Krol, J. P. Ebel, E. Lazar, B. Haendler, and M. Jacob, "U2 RNA shares a structural domain with U1, U4, and U5 RNAs," *EMBO Journal*, vol. 1, no. 10, pp. 1259–1263, 1982.

[18] H. Shibata, T. S. Ro Choi, R. Reddy, Y. C. Choi, D. Henning, and H. Bush, "The primary nucleotide sequence of nuclear U2 ribonucleic acid. The 5' terminal portion of the molecule," *Journal of Biological Chemistry*, vol. 250, no. 10, pp. 3909–3920, 1975.

[19] D. Suh, H. Busch, and R. Reddy, "Isolation and characterization of a human U3 small nucleolar RNA gene," *Biochemical and Biophysical Research Communications*, vol. 137, no. 3, pp. 1133–1140, 1986.

[20] K. A. Parker and J. A. Steitz, "Structural analysis of the human U3 ribonucleoprotein particle reveal a conserved sequence available for base pairing with pre-rRNA," *Molecular and Cellular Biology*, vol. 7, no. 8, pp. 2899–2913, 1987.

[21] T. S. Ro-Choi, "Nucleolar snoRNA and ribosome production," *Molecules and Cells*, vol. 7, no. 4, pp. 451–467, 1997.

[22] A. Krol, C. Branlant, E. Lazar, H. Gallinaro, and M. Jacob, "Primary and secondary structures of chicken, rat and man nuclear U4 RNAs. Homologies with U1 and U5 RNAs," *Nucleic Acids Research*, vol. 9, no. 12, pp. 2699–2716, 1981.

[23] S. Liu, P. Li, O. Dybkov et al., "Binding of the human Prp31 Nop domain to a composite RNA-protein platform in U4 snRNP," *Science*, vol. 316, no. 5821, pp. 115–120, 2007.

[24] R. Reddy, D. Henning, and H. Busch, "The primary nucleotide sequence of U4 RNA," *Journal of Biological Chemistry*, vol. 256, no. 7, pp. 3532–3538, 1981.

[25] A. Krol, H. Gallinaro, E. Lazar et al., "The nuclear 5S RNAs from chicken, rat and man. U5 RNAs are encoded by multiple genes," *Nucleic Acids Research*, vol. 9, no. 4, pp. 769–787, 1981.

[26] R. Reddy and H. Busch, "U-snRNAs of nuclear RNPs," in *The Cell Nucleus*, H. Busch, Ed., vol. 8, chapter 7, pp. 261–306, Academic Press, New York, NY, USA, 1981.

[27] G. R. Kunkel, R. L. Maser, J. P. Calvet, and T. Pederson, "U6 small nuclear RNA is transcribed by RNA polymerase III," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 83, no. 22, pp. 8575–8579, 1986.

[28] P. Epstein, R. Reddy, D. Henning, and H. Busch, "The nucleotide sequence of nuclear U6 (4.7 S) RNA," *Journal of Biological Chemistry*, vol. 255, no. 18, pp. 8901–8906, 1980.

[29] K. L. Mowry and J. A. Steitz, "Identification of the human U7 snRNP as one of several factors involved in the 3' end maturation of histone premessenger RNA's," *Science*, vol. 238, no. 4834, pp. 1682–1687, 1987.

[30] K. Tyc and J. A. Steitz, "U3, U8 and U13 comprise a new class of mammalian snRNPs localized in the cell nucleolus," *EMBO Journal*, vol. 8, no. 10, pp. 3113–3119, 1989.

[31] T. S. Ro, M. Muramatsu, and H. Busch, "Labeling of RNA of isolated nucleoli with UTP-2-^{14}C," *Biochemical and Biophysical Research Communications*, vol. 14, no. 2, pp. 149–155, 1964.

[32] Y. C. Choi, K. Smetana, and H. Busch, "Studies on RNA isolated from nucleolar fractions of Novikoff hepatoma cells by sequential extractions with dilute and concentrated salt solutions," *Experimental Cell Research*, vol. 53, no. 2-3, pp. 582–602, 1968.

[33] E. S. Maxwell and M. J. Fournier, "The small nucleolar RNAs," *Annual Review of Biochemistry*, vol. 64, pp. 897–934, 1995.

[34] K. T. Tycowski, N.G. Kolev, N. K. Conrad, V. Fok, and J. A. Steitz, "The ever-growing world of small nuclear ribonucleoproteins," in *The RNA World*, R. F. Gesteland, T. R. Cech, and J. F. Atkins, Eds., chapter 12, pp. 327–368, CSHL Press, 2006.

[35] C. L. Will and R. Lührmann, "Spliceosome structure and function," in *The RNA World*, R. F. Gesteland, T. R. Cech, and J. F. Atkins, Eds., chapter 13, pp. 369–400, CSHL Press, 2006.

[36] T. S. Ro-Choi, R. Reddy, Y. C. Choi, N. B. Raj, and D. Henning, "Primary sequence of U1 nuclear RNA and unusual feature of 5' end structure of LMWN RNA," *Proceedings Federation of American Societies for Experimental Biology*, vol. 33, p. 1548, 1974.

[37] T. S. Ro-Choi, "Nuclear snRNA and nuclear function (discovery of 5' cap structures in RNA)," *Critical Reviews in Eukaryotic Gene Expression*, vol. 9, no. 2, pp. 107–158, 1999.

[38] T. S. Ro-Choi, Y. Moriyama, Y. C. Choi, and H. Busch, "Isolation and purification of a nuclear 4.5S ribonucleic acid of the Novikoff hepatoma," *Journal of Biological Chemistry*, vol. 245, no. 8, pp. 1970–1977, 1970.

[39] S. M. el-Khatib, T. S. Ro-Choi, Y. C. Choi, and H. Busch, "Studies on nuclear 4.5 S ribonucleic acid III of Novikoff hepatoma ascites cells," *Journal of Biological Chemistry*, vol. 245, no. 13, pp. 3416–3421, 1970.

[40] G. Dreyfuss, M. J. Matunis, S. Piñol-Roma, and C. G. Burd, "hnRNP proteins and the biogenesis of mRNA," *Annual Review of Biochemistry*, vol. 62, pp. 289–321, 1993.

[41] J. C. Long and J. F. Caceres, "The SR protein family of splicing factors: master regulators of gene expression," *Biochemical Journal*, vol. 417, no. 1, pp. 15–27, 2009.

[42] G. Dreyfuss, "Structure and function of nuclear and cytoplasmic ribonucleoprotein particles," *Annual Review of Cell Biology*, vol. 2, pp. 459–498, 1986.

[43] W. M. LeStourgeon, S. F. Barnett, and S. J. Northington, "Tetramer of core proteins of 40S nuclear ribonucleoprotein particles assemble to package nascent transcripts into a repeating array of regular particles," in *The Eukaryotic Nucleus*, P. R. Strauss and S. H. Willson, Eds., vol. 2, pp. 477–502, The Telford Press, NJ, USA, 1990.

[44] M. Zuker, "Computer prediction of RNA structure," in *Methods in Enzymology*, vol. 180, chapter 20, pp. 262–288, 1989.

[45] D. E. Draper, "Protein-RNA recognition," *Annual Review of Biochemistry*, vol. 64, pp. 593–620, 1995.

[46] H.-J. Kim and K. Han, "Automated modeling of the RNA folding process," *Molecules and Cells*, vol. 5, pp. 406–412, 1995.

[47] T. S. Ro-Choi and Y. C. Choi, "Thermodynamic analyses of the constitutive splicing pathway for ovomucoid pre-mRNA," *Molecules and Cells*, vol. 27, no. 6, pp. 657–665, 2009.

[48] T. S. Ro-Choi and Y. C. Choi, "Structural elements of dynamic RNA strings," *Molecules and Cells*, vol. 16, no. 2, pp. 201–210, 2003.

[49] T. S. Ro-Choi and Y. C. Choi, "A modeling study of Co-transcriptional metabolism of hnRNP using FMR1 Gene," *Molecules and Cells*, vol. 23, no. 2, pp. 228–238, 2007.

[50] J. S. Mattick, "The functional genomics of noncoding RNA," *Science*, vol. 309, no. 5740, pp. 1527–1528, 2005.

[51] J. M. Claverie, "Fewer genes, more noncoding RNA," *Science*, vol. 309, no. 5740, pp. 1529–1530, 2005.

[52] U. A. Ørom, T. Derrien, M. Beringer et al., "Long noncoding RNAs with enhancer-like function in human cells," *Cell*, vol. 143, no. 1, pp. 46–58, 2010.

[53] W. J. Steele, N. Okamura, and H. Busch, "Effects of thioacetamide on the composition and biosynthesis of nucleolar and nuclear ribonucleic acid in rat liver," *Journal of Biological Chemistry*, vol. 240, pp. 1742–1749, 1965.

[54] R. W. Holley, G. A. Everett, J. T. Madison, and A. Zamir, "Nucleotide sequence in the yeast alanine transfer ribonucleic acid," *Journal of Biological Chemistry*, vol. 240, pp. 2122–2128, 1965.

[55] R. de Wachter and W. Fiers, "Studies on the bacteriophage MS2. IV. The 3′-OH terminal undecanucleotide sequence of the viral RNA chain," *Journal of Molecular Biology*, vol. 30, no. 3, pp. 507–518, 1967.

[56] M. Silberklang, A. M. Gillum, and U. L. RajBhandary, "[3] Use of in Vitro ^{32}P labeling in the sequence analysis of nonradioactive tRNAs," *Methods in Enzymology*, vol. 59, no. C, pp. 58–109, 1979.

[57] W. Y. Li, R. Reddy, D. Henning, P. Epstein, and H. Busch, "Nucleotide sequence of 7 S RNA. Homology to Alu DNA and La 4.5 S RNA," *Journal of Biological Chemistry*, vol. 257, no. 9, pp. 5136–5142, 1982.

[58] T. S. Ro-Choi, "U3 RNA cap structure, $m_3^{2,2,7}$GpppAmpA(m)pApG, of Novikoff hepatoma cell," *Molecules and Cells*, vol. 6, no. 4, pp. 436–443, 1996.

[59] P. Carninci and Y. Hayashizaki, "High-efficiency full-length cDNA cloning," *Methods in Enzymology*, vol. 303, pp. 19–44, 1999.

[60] D. A. Peattie, "Direct chemical method for sequencing RNA," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 76, no. 4, pp. 1760–1764, 1979.

[61] Y. Kuchino and S. Nishimura, "Enzymatic RNA sequencing," *Methods in Enzymology*, vol. 180, pp. 154–163, 1989.

[62] B. G. Barrel, "Fractionation and sequence analysis of radioactive nucleotides," *Procedures in Nucleic Acid Research*, vol. 2, pp. 751–779, 1971.

[63] K. Randerath, R. C. Gupta, and E. Randerath, "[^{3}H] and [^{32}P] derivative methods for base composition and sequence analysis of RNA," *Methods in Enzymology*, vol. 65, no. C, pp. 638–680, 1971.

[64] Y. C. Choi and H. Busch, "Modified nucleotides in T1 RNase oligonucleotides of 18S ribosomal RNA of the Novikoff hepatoma," *Biochemistry*, vol. 17, no. 13, pp. 2551–2560, 1978.

[65] International Human Genome Sequencing Consortium, "Initial sequencing and analysis of the human genome," *Nature*, vol. 409, no. 6838, pp. 860–921, 2001.

[66] J. Craig Venter, M. D. Adams, E. W. Myers et al., "The sequence of the human genome," *Science*, vol. 291, no. 5507, pp. 1304–1351, 2001.

[67] R. H. Waterston, K. Lindblad-Toh, E. Birney et al., "Initial sequencing and comparative analysis of the mouse genome," *Nature*, vol. 420, no. 6915, pp. 520–562, 2002.

[68] G. S. Wang and T. A. Cooper, "Splicing in disease: disruption of the splicing code and the decoding machinery," *Nature Reviews Genetics*, vol. 8, no. 10, pp. 749–761, 2007.

[69] K. J. Hertel, "Combinatorial control of exon recognition," *Journal of Biological Chemistry*, vol. 283, no. 3, pp. 1211–1215, 2008.

[70] S. Hausmann, S. Zheng, M. Costanzo et al., "Genetic and biochemical analysis of yeast and human cap trimethylguanosine synthase: functional overlap of 2,2,7-trimethylguanosine caps, small nuclear ribonucleoprotein components, PRE-mRNA splicing factors, and RNA decay pathways," *Journal of Biological Chemistry*, vol. 283, no. 46, pp. 31706–31718, 2008.

[71] P. Misra, C. Qi, S. Yu et al., "Interaction of PIMT with transcriptional coactivators CBP, p300, and PBP differential role in transcriptional regulation," *Journal of Biological Chemistry*, vol. 277, no. 22, pp. 20011–20019, 2002.

[72] J. Mouaikel, U. Narayanan, C. Verheggen et al., "Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron," *EMBO Reports*, vol. 4, no. 6, pp. 616–622, 2003.

[73] P. T. Gilham, "An addition reaction specific for uridine and guanosine nucleotides and its application to the modification of ribonuclease action," *Journal of the American Chemical Society*, vol. 84, no. 4, pp. 687–688, 1962.

[74] R. Naylor, N. W. Y. Ho, and P. T. Gilham, "Selective chemical modifications of uridine and pseudouridine in polynucleotides and their effect on the specificities of ribonuclease and phosphodiesterases," *Journal of the American Chemical Society*, vol. 87, no. 18, pp. 4209–4210, 1965.

[75] N. W. Y. Ho and P. T. Gilham, "Reaction of pseudou-ridine and inosine with N-cyclohexyl-N′-β-(4-methylmorpholinium) ethylcarbodiimide," *Biochemistry*, vol. 10, no. 20, pp. 3651–3657, 1971.

[76] A. G. Saponara and M. D. Enger, "Occurrence of N2,N2,7-methylguanosine in minor RNA species of a mammalian cell line," *Nature*, vol. 223, no. 5213, pp. 1365–1366, 1969.

[77] T. S. Ro-Choi, R. Redy, D. Henning, T. Takano, C. W. Taylor, and H. Busch, "Nucleotide sequence of 4.5S Ribonucleic Acid I of Novikoff hepatoma cell nuclei," *Journal of Biological Chemistry*, vol. 247, no. 10, pp. 3205–3222, 1972.

[78] S. Schwartz, N. Gal-Mark, N. Kfir, R. Oren, E. Kim, and G. Ast, "Alu exonization events reveal features required for precise recognition of exons by the splicing machinery," *PLoS Computational Biology*, vol. 5, no. 3, Article ID e1000300, 2009.

[79] K. Nagai, C. Oubridge, A. Kuglstatter, E. Menichelli, C. Isel, and L. Jovine, "Structure, function and evolution of the signal recognition particle," *EMBO Journal*, vol. 22, no. 14, pp. 3479–3485, 2003.

[80] A. W. Prestayko, M. Tonato, B. C. Lewis, and H. Busch, "Heterogeneity of nucleolar U3 ribonucleic acid of the Novikoff hepatoma," *Journal of Biological Chemistry*, vol. 246, no. 1, pp. 182–187, 1971.

[81] E. Randerath, C. T. Yu, and K. Randerath, "Base analysis of ribopolynucleotides by chemical tritium labeling: a methodological study with model nucleosides and purified tRNA species," *Analytical Biochemistry*, vol. 48, no. 1, pp. 172–198, 1972.

[82] R. Reddy, T. S. Ro-Choi, D. Henning, H. Shibata, Y. C. Choi, and H. Busch, "Modified nucleosides of nuclear and nucleolar low molecular weight ribonucleic acid," *Journal of Biological Chemistry*, vol. 247, no. 22, pp. 7245–7250, 1972.

[83] C. M. Mauritzen, Y. C. Choi, and H. Busch, "Preparation of Macromolecules of Very High Specific Activity of Tumor cells," in *Methods in Cancer Research*, H. Busch, Ed., vol. 6, pp. 253–282, Academic Press, New York, NY, USA, 1970.

[84] Y. C. Choi and T. S. Ro-Choi, "Basic characteristics of different classes of cellular RNAs; a directory," in *Cell Biology; A Comprehensive Treatise*, L. Goldstein and D. Prescott, Eds., vol. 3, chapter 12, pp. 609–667, New York Academic, New York, NY, USA, 1980.

[85] H. Miyoshi, D. S. Dwyer, B. D. Keiper, M. Jankowska-Anyszka, E. Darzynkiewicz, and R. E. Rhoads, "Discrimination between mono- and trimethylated cap structures by two isoforms of Caenorhabditis elegans eIF4E," *EMBO Journal*, vol. 21, no. 17, pp. 4680–4690, 2002.

[86] C. Mazza, M. Ohno, A. Segref, I. W. Mattaj, and S. Cusack, "Crystal structure of the human nuclear cap binding complex," *Molecular Cell*, vol. 8, no. 2, pp. 383–396, 2001.

[87] C. Mazza, A. Segref, I. W. Mattaj, and S. Cusack, "Large-scale induced fit recognition of an m^7GpppG cap analogue by the human nuclear cap-binding complex," *EMBO Journal*, vol. 21, no. 20, pp. 5548–5557, 2002.

[88] C. M. Smith and J. A. Steitz, "Sno storm in the nucleolus: new roles for myriad small RNPs," *Cell*, vol. 89, no. 5, pp. 669–672, 1997.

[89] C. Y. Lee, A. Lee, and G. Chanfreau, "The roles of endonucleolytic cleavage and exonucleolytic digestion in the 5′-end processing of S. cerevisiae box C/D snoRNAs," *RNA*, vol. 9, no. 11, pp. 1362–1370, 2003.

[90] T. Monecke, A. Dickmanns, and R. Ficner, "Structural basis for m^7G-cap hypermethylation of small nuclear, small nucleolar and telomerase RNA by the dimethyltransferase TGS1," *Nucleic Acids Research*, vol. 37, no. 12, pp. 3865–3877, 2009.

[91] A. Strasser, A. Dickmanns, R. Lührmann, and R. Ficner, "Structural basis for m_3G-cap-mediated nuclear import of spliceosomal UsnRNPs by snurportin1," *EMBO Journal*, vol. 24, no. 13, pp. 2235–2243, 2005.

[92] J. L. Chen, M. A. Blasco, and C. W. Greider, "Secondary structure of vertebrate telomerase RNA," *Cell*, vol. 100, no. 5, pp. 503–514, 2000.

[93] J. Franke, J. Gehlen, and A. E. Ehrenhofer-Murray, "Hypermethylation of yeast telomerase RNA by the snRNA and snoRNA methyltransferase Tgs1," *Journal of Cell Science*, vol. 121, no. 21, pp. 3553–3560, 2008.

[94] B. E. Jády, E. Bertrand, and T. Kiss, "Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal," *Journal of Cell Biology*, vol. 164, no. 5, pp. 647–652, 2004.

[95] S. Bektesh, K. Van Doren, and D. Hirsh, "Presence of the Caenorhabditis elegans spliced leader on different mRNAs and in different genera of nematodes," *Genes & Development*, vol. 2, no. 10, pp. 1277–1283, 1988.

[96] T. Blumenthal, "Trans-splicing and operons," *WormBook*, pp. 1–9, 2005.

[97] C. K. Ho and S. Shuman, "Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme," *Molecular Cell*, vol. 3, no. 3, pp. 405–411, 1999.

[98] K. Mizumoto and Y. Kaziro, "Messenger RNA capping enzymes from eukaryotic cells," *Progress in Nucleic Acid Research and Molecular Biology*, vol. 34, no. C, pp. 1–28, 1987.

[99] D. Testa and A. K. Banerjee, "Two methyltransferase activities in the purified virions of vesicular stomatitis virus," *Journal of Virology*, vol. 24, no. 3, pp. 786–793, 1977.

[100] A. J. Shatkin, "Capping of eucaryotic mRNAs," *Cell*, vol. 9, no. 4, pp. 645–653, 1976.

[101] S. R. Langberg and B. Moss, "Post-transcriptional modifications of mRNA. Purification and characterization of cap I and cap II RNA (nucleoside-2'-)-methyltransferases from HeLa cells," *Journal of Biological Chemistry*, vol. 256, no. 19, pp. 10054–10060, 1981.

[102] D. L. Spector, "Colocalization of U1 and U2 small nuclear RNPs by immunocytochemistry," *Biology of the Cell*, vol. 51, no. 1, pp. 109–112, 1984.

[103] G. Plessel, U. Fischer, and R. Luhrmann, "m_3G cap hypermethylation of U1 small nuclear ribonucleoprotein (snRNP) in vitro: evidence that the U1 small nuclear RNA-(guanosine-N2)- methyltransferase is a non-snRNP cytoplasmic protein that requires a binding site on the Sm core domain," *Molecular and Cellular Biology*, vol. 14, no. 6, pp. 4160–4172, 1994.

[104] I. W. Mattaj, "Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding," *Cell*, vol. 46, no. 6, pp. 905–911, 1986.

[105] C. Verheggen, D. L. J. Lafontaine, D. Samarsky et al., "Mammalian and yeast U3 snoRNPs are matured in specific and related nuclear compartments," *EMBO Journal*, vol. 21, no. 11, pp. 2736–2745, 2002.

[106] J. Mouaikel, C. Verheggen, E. Bertrand, J. Tazi, and R. Bordonné, "Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus," *Molecular Cell*, vol. 9, no. 4, pp. 891–901, 2002.

[107] S. J. Baserga, M. Gilmore-Hebert, and X. W. Yang, "Distinct molecular signals for nuclear import of the nucleolar snRNA, U3," *Genes and Development*, vol. 6, no. 6, pp. 1120–1130, 1992.

[108] M. P. Terns and J. E. Dahlberg, "Retention and 5' cap trimethylation of U3 snRNA in the nucleus," *Science*, vol. 264, no. 5161, pp. 959–961, 1994.

[109] Y. Zhu, C. Qi, W. Q. Cao, A. V. Yeldandi, M. S. Rao, and J. K. Reddy, "Cloning and characterization of PIMT, a protein with a methyltransferase domain, which interacts with and enhances nuclear receptor coactivator PRIP function," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 18, pp. 10380–10385, 2001.

[110] C. Girard, C. Verheggen, H. Neel et al., "Characterization of a short isoform of human Tgs1 hypermethylase associating with small nucleolar ribonucleoprotein core proteins and produced by limited proteolytic processing," *Journal of Biological Chemistry*, vol. 283, no. 4, pp. 2060–2069, 2008.

[111] O. Komonyi, G. Pápai, I. Enunlu et al., "DTL, the Drosophila homolog of PIMT/Tgs1 nuclear receptor coactivator-interacting protein/RNA methyltransferase, has an essential role in development," *Journal of Biological Chemistry*, vol. 280, no. 13, pp. 12397–12404, 2005.

[112] S. Hausmann and S. Shuman, "Specificity and mechanism of RNA cap guanine-N2 methyltransferase (Tgs1)," *Journal of Biological Chemistry*, vol. 280, no. 6, pp. 4021–4024, 2005.

[113] S. Hausmann and S. Shuman, "Giardia lamblia RNA cap guanine-N2 methyltransferase (Tgs2)," *Journal of Biological Chemistry*, vol. 280, no. 37, pp. 32101–32106, 2005.

[114] J. D. Bangs, P. F. Crain, T. Hashizume, J. A. McCloskey, and J. C. Boothroyd, "Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides," *Journal of Biological Chemistry*, vol. 267, no. 14, pp. 9805–9815, 1992.

[115] B. Mittra, J. R. Zamudio, J. M. Bujnicki et al., "The TbMTr1 spliced leader RNA cap 1 2′-O-ribose methyltransferase from Trypanosoma brucei acts with substrate specificity," *Journal of Biological Chemistry*, vol. 283, no. 6, pp. 3161–3172, 2008.

[116] Y. Takagi, S. Sindkar, D. Ekonomidis, M. P. Hall, and C. K. Ho, "Trypanosoma brucei encodes a bifunctional capping enzyme essential for cap 4 formation on the spliced leader RNA," *Journal of Biological Chemistry*, vol. 282, no. 22, pp. 15995–16005, 2007.

[117] M. P. Hall and C. Kiong Ho, "Characterization of a Trypanosoma brucei RNA cap (guanine N-7) methyltransferase," *RNA*, vol. 12, no. 3, pp. 488–497, 2006.

[118] E. Silva, E. Ullu, R. Kobayashi, and C. Tschudi, "Trypanosome capping enzymes display a novel two-domain structure," *Molecular and Cellular Biology*, vol. 18, no. 8, pp. 4612–4619, 1998.

[119] J. A. McCloskey, "Constituents of nucleic acids: overview and strategy," *Methods in Enzymology*, vol. 193, pp. 771–790, 1990.

[120] R. J. Taft, K. C. Pang, T. R. Mercer, M. Dinger, and J. S. Mattick, "Non-coding RNAs: regulators of disease," *Journal of Pathology*, vol. 220, no. 2, pp. 126–139, 2010.

SAGE-Hindawi Access to Research
Journal of Nucleic Acids
Volume 2011, Article ID 748632, 17 pages
doi:10.4061/2011/748632

Research Article

Site-Selective Artificial Ribonucleases: Oligonucleotide Conjugates Containing Multiple Imidazole Residues in the Catalytic Domain

Natalia G. Beloglazova,[1] Martin M. Fabani,[2] Nikolai N. Polushin,[3] Vladimir V. Sil'nikov,[1] Valentin V. Vlassov,[1] Elena V. Bichenkova,[2] and Marina A. Zenkova[1]

[1] *Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences, Novosibirsk 630090, Russia*
[2] *School of Pharmacy and Pharmaceutical Sciences, The University of Manchester, Manchester M13 9PL, UK*
[3] *Fidelity Systems Inc., 7961 Cessna Avenue, Gaithersburg, MD 20879, USA*

Correspondence should be addressed to Marina A. Zenkova, marzen@niboch.nsc.ru

Received 7 April 2011; Accepted 5 July 2011

Academic Editor: Dmitry A. Stetsenko

Design of site-selective artificial ribonucleases (aRNases) is one of the most challenging tasks in RNA targeting. Here, we designed and studied oligonucleotide-based aRNases containing *multiple* imidazole residues in the catalytic part and systematically varied structure of cleaving constructs. We demonstrated that the ribonuclease activity of the conjugates is strongly affected by the number of imidazole residues in the catalytic part, the length of a linker between the catalytic imidazole groups of the construct and the oligonucleotide, and the type of anchor group, connecting linker structure and the oligonucleotide. Molecular modeling of the most active aRNases showed that preferable orientation(s) of cleaving constructs strongly depend on the structure of the anchor group and length of the linker. The inclusion of deoxyribothymidine anchor group significantly reduced the probability of cleaving groups to locate near the cleavage site, presumably due to a stacking interaction with the neighbouring nucleotide residue. Altogether the obtained results show that dynamics factors play an important role in site-specific RNA cleavage. Remarkably high cleavage activity was displayed by the conjugates with the most flexible and extended cleaving construct, which presumably provides a better opportunity for imidazole residues to be correctly positioned in the vicinity of scissile phosphodiester bond.

1. Introduction

The idea of site-selective artificial ribonucleases which are capable of cleaving any particular RNA sequence *in vitro* and *in vivo*, is a very attractive approach as, apart from being useful tools in molecular biology, these chemical ribonucleases are anticipated to be helpful for target validation, and even for the development of potential antiviral or anticancer therapeutics. The idealized site-selective artificial ribonuclease would be a compound that is easily synthesized, chemically stable, targeted to any chosen RNA sequence, and highly efficient in cleaving of the phosphodiester bonds. The oligonucleotide-based artificial ribonucleases seem to meet the majority of these criteria due to the fact that they can be directed to almost any desired RNA region in a site specific-manner by controlling the sequence of the oligonucleotide recognition part.

Virtually all types of the reactive groups, which are known to catalyze RNA scission, have been exploited for design and preparation of artificial ribonucleases [1–3]. Conjugates of oligonucleotides and RNA cleaving groups can be synthesized using two main approaches: incorporation of a catalytic structure in the oligonucleotide in the course of standard synthesis or by postsynthetic derivatization of an unprotected oligonucleotide [4]. Several alternative procedures for each strategy are available. The first approach often involves the automated oligonucleotide synthesis, when phosphoramidites of suitably protected nucleosides and

catalytic groups are used as building blocks. Monomeric building blocks containing various metal complexes [5–8] or imidazole-based structures [9–16] have been previously proposed for the synthesis of the oligonucleotide-based artificial ribonucleases. The successive synthesis of oligonucleotide and catalytic structures of conjugates on the same solid-phase support is a promising variation of this approach [17–19]. However, this approach is inappropriate for the majority of the potential users of the conjugates due to the necessity to synthesize the unique phosphoramidite derivatives of the corresponding catalytic structures in each particular case, which creates the problems associated with the choice of appropriate protection groups and the deprotection procedure.

Various RNA-cleaving catalysts have been tethered to the 5′-terminus of oligonucleotides. However, the available synthetic methods, except for a few examples [20], allow the incorporation of only a single functional moiety per phosphoramidite unit. Generally, 5′-tethering can be implemented in two ways either by direct incorporation of a suitably protected and activated constructs into the 5′-terminus at the end of standard solid phase synthesis or by postsynthetic derivatization. In the former case, synthesis of highly modified oligonucleotides may become somewhat problematic because increase in a number of incorporations is expected to substantially decrease the overall yield of the conjugated oligonucleotides. Alternatively, a precursor molecule can first be formed by reacting the 5′-OH group of an assembled oligonucleotide with a heterobifunctional reagent (first modifier) bearing both the phosphoramidite moiety and a reactive group. The reactive group of the precursor oligonucleotide is then postsynthetically derivatized with an appropriate functional additive (second modifier) [21, 22]. This precursor strategy has the obvious advantage that the same parent compound can be used to synthesize a number of differently tethered products. The first generation of oligonucleotide-based artificial ribonucleases, which was prepared using the precursor technique, has recently been reported to display high cleaving activity [23, 24]. The precursor approach seems to be advantageous for the synthesis of libraries of oligonucleotide conjugates containing different functional groups.

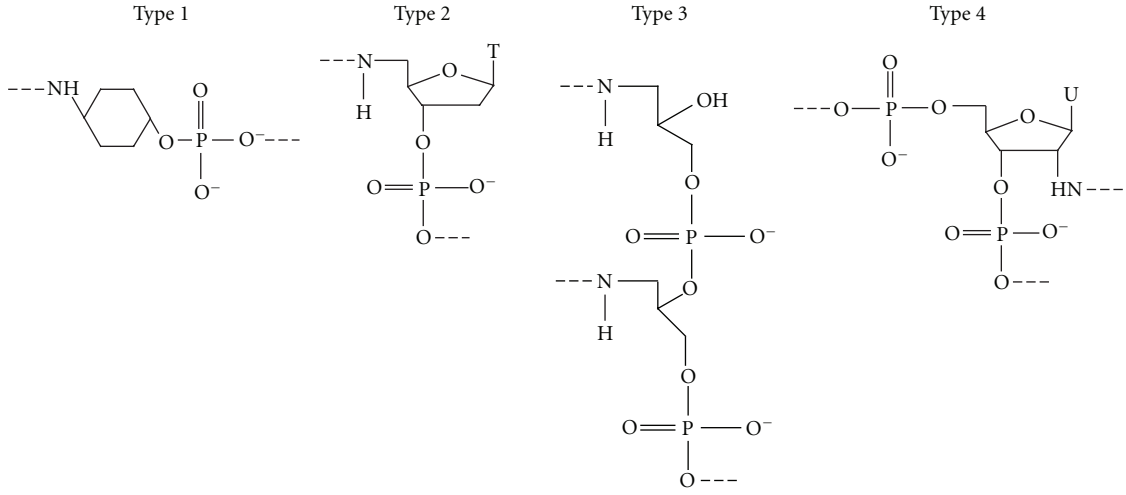
However, the cleavage activity of these artificial ribonucleases so far has been inferior to the natural counterparts in terms of targeting and/or cleavage efficiency. We anticipate that the high level of efficiency of RNA cleavage by these oligonucleotide-based artificial ribonucleases can only be achieved if the optimal arrangement of all groups involved in the binding and catalysis is accomplished. It might be possible to improve on this by iterative design of the binding and cleavage domains by taking into account the optimal structures of the cleaving groups and their mutual spatial orientations against the RNA target. The structural organisation of artificial ribonucleases and the dynamic behavior of their cleaving constructs seem to be the key factors contributing to the cleavage activity of these compounds.

Several years ago, we reported a new type of oligonucleotide-based aRNases containing *bis*-imidazole cleaving constructs, which were prepared by precursor technique. It was demonstrated that these highly active *bis*-imidazole cleaving constructs were extremely flexible and the cleavage of the phosphodiester bond seemed to be a random event, which happened with the high efficiency every time, when these catalytic groups approached the scissile phosphodiester bond [23, 24]. In the present paper, we use the developed precursor approach to synthesize the scaffolds of aRNases with the systematically varied structure of the cleaving constructs bearing multiple imidazole residues in the catalytic part. We demonstrate here that some of these conjugates exhibit remarkably high cleavage activity within the $tRNA^{Phe}$ *conjugate* hybrid and identify possible location(s) of cleaving group(s) relative to the scissile $^{5'}$C63-A64 site. We also show, here, that the cleaving activity of the conjugates is governed by the intrinsic properties of the catalytic constructs, but not the detailed structure of the DNA:RNA hybrid, which was the same in all experiments.

2. Materials and Methods

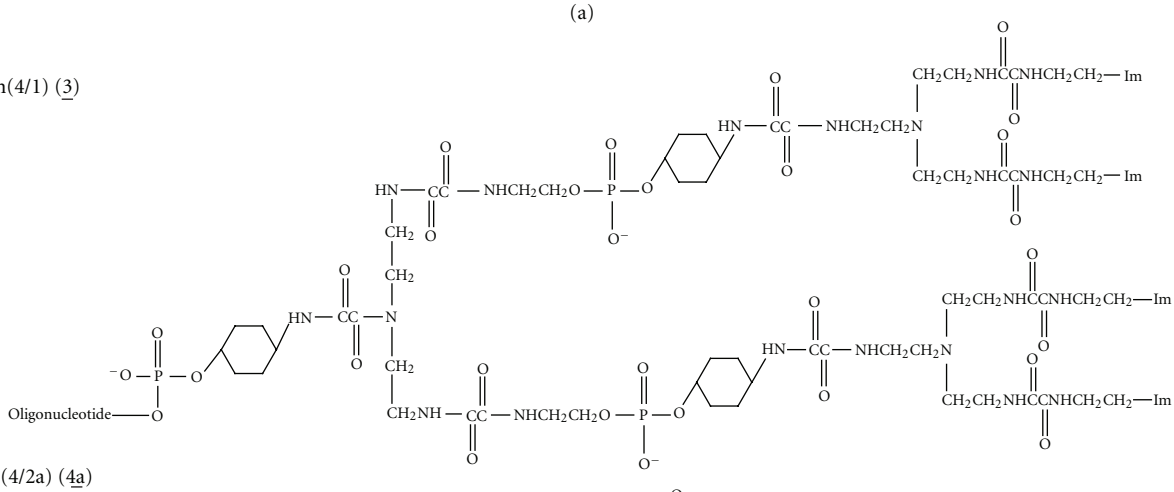
2.1. General Chemicals. All buffers used in the experiments were prepared using milliQ water, contained 0.1 mM EDTA, and were filtered through filters with pore size 0.22 μm (Millipore). [5′-^{32}P]-pCp was from Biosan Co., Russia. T4 RNA ligase was purchased from Fermentas (*Lethuania*). Oligonucleotides TGGTGCGAATTCTG (**A**) and GATCGAACACAGGACCT (**B**) were synthesized on ASM-700 DNA Synthesizer (Biosset, Russia) by standard solid-phase phosphoramidite procedure. Yeast $tRNA^{Phe}$ was a generous gift from Dr. G. Keith (Institut de Biologie Moléculaire et Cellulaire du CNRS, Strasbourg, France). Methoxyoxalamido (MOX) modifiers **M** used for conjugate synthesis were prepared as described in [21, 24].

2.2. Synthesis of Oligonucleotide Conjugates B-Im(n) Type 1, Type 2, Type 3, and Type 4. Deoxyribooligonucleotide **B** (GATCGAACACAGGACCT) was synthesized using standard solid-phase phosphoramidite chemistry, except that the dC^{Bz}-phosphoramidite was replaced with the dC^{Ac}-phosphoramidite. At the end of the automatic synthesis, the core 17-mer was coupled with one of the modifiers **M** (Type 1, Type 2, Type 3, or Type 4; see Figure 1) in 0.2 M in acetonitrile for 15 min. The prepared oligonucleotide MOX precursors were then functionalized with 2 M solution of histamine in dimethylformamide for 3 hrs at 20°C at constant shaking followed by deprotection with ammonia to produce the oligonucleotide conjugates containing multiple-imidazole catalytic structures. The conjugates were purified by electrophoresis in 16% polyacrylamide/8 M urea gel under denaturing conditions. The purified conjugates were analyzed by ESI and/or MALDI MS (calculated and experimental values of molecular masses are listed in Figure 1). In the cases when MALDI-MS analysis did not produce the expected mass spectra due to the conjugate fragmentation upon analysis, the conjugates were analyzed by electrophoresis under denaturing conditions (Figure 1(c)) and by HPLC (primary data not shown). Conjugates were designated as **B-Im(N/m)**, where B-Im corresponds to the oligonucleotide B and shows that Im are in catalytic part, N is the number of

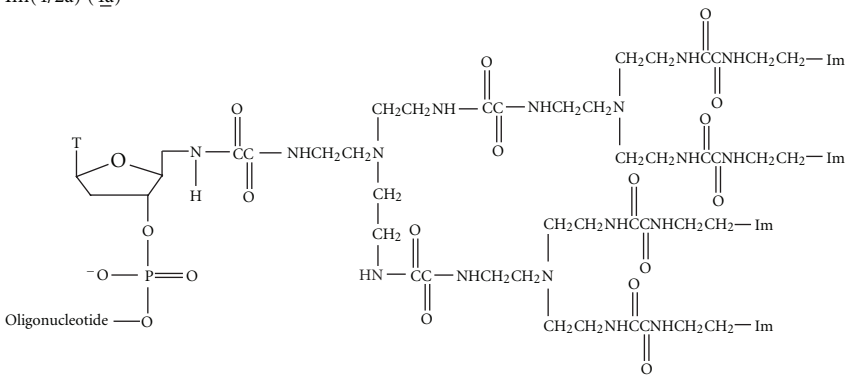


(a)

B-Im(4/1) (3)



B-Im(4/2a) (4a)



B-Im(4/2b) (4b)

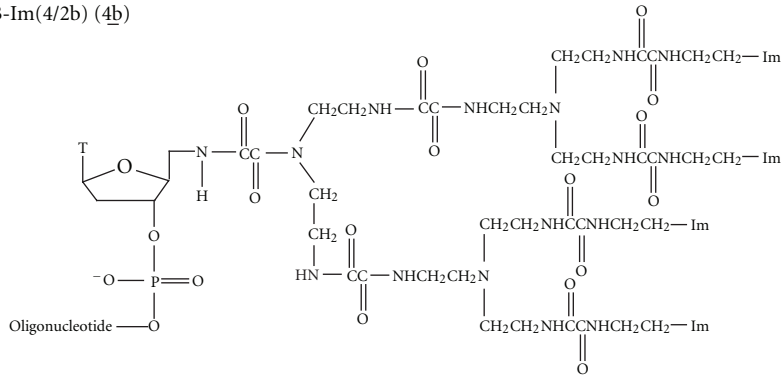


Figure 1: Continued.

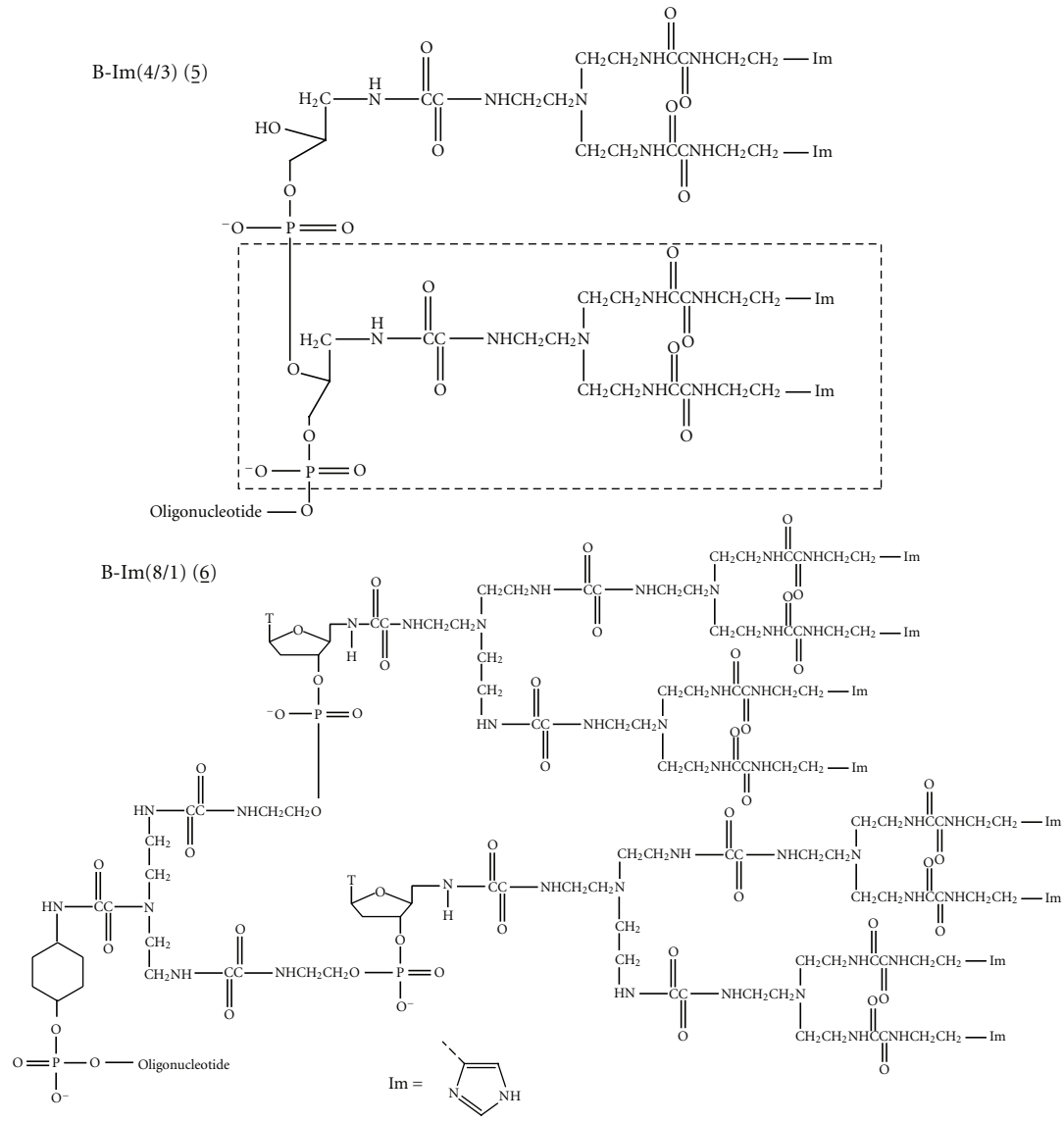


Figure 1: *Schematic representation of the catalytic structures of the site-selective artificial ribonucleases B-Im(N/m) and analysis of the conjugates.* (a) Anchor groups used for attachment of dendrimeric RNA-cleaving constructs to oligonucleotide **B** 5′-phosphate. (b) Schematic representation of conjugates bearing four and eight imidazole residues and anchor groups of Type1, Type 2, and Type 3, obtained using precursor technology. Mass spectrometry data: **1**: calculated mass 5523.30, found mass 5523.73; **2**: 5566.80/5566.39; **3**: 5692.89/5692.03; **4a:** 6838.95/6838,65; **5**: 6423.63/6425.03. (c) Analysis of the conjugates **B-Im(N/m)** by 12% polyacrylamide/8 M urea gel electrophoresis. 0.1 A_{260} unit of each conjugate was applied on the gel; the conjugates were visualized in the gel by staining with Stains-All. Numbers on the top of the gel correspond to numeration of the conjugates in the Table 1 and Figure 1(b) and Supplementary Material Figure 1.

imidazole residues (from 2 to 32), and m corresponds to the type of anchor group (1, 2, 3, or 4; see Figure 1(a)).

2.3. Cleavage of tRNA^{Phe} with Oligonucleotide Conjugates. [3′-^{32}P]-tRNA^{Phe} was prepared and purified as described in [24]. 3′-end labeled tRNA was dissolved in water and stored at −20°C. Specific radioactivity of the [^{32}P]-tRNA^{Phe} was 5×10^5 cpm/pmole.

Standard reaction mixture (10 μL) contained imidazole buffer (50 mM imidazole buffer, pH 7.0, 200 mM KCl, 0.1 mM EDTA, 100 μg/mL total tRNA from *Escherichia coli* as carrier), 5×10^{-7} M [3′-^{32}P]-tRNA^{Phe}, and one of oligonucleotide conjugates **B-Im(N/m)** at concentrations ranging from 5×10^{-7} to 5×10^{-4} M (as indicated in the legends in the figures). Reactions were carried out at 37°C and were quenched by precipitation of tRNA and tRNA fragments with 150 μL of 2% lithium perchlorate solution in acetone. RNA was collected by centrifugation and dissolved in loading buffer (6 M urea, 0.025% bromophenol blue, 0.025% xylene cyanol). tRNA cleavage products were analyzed by electrophoresis in 12% polyacrylamide/8 M urea gel. To identify cleavage sites, an imidazole ladder [24] and a G-ladder produced by partial tRNA cleavage with RNase T1 [25] were run in parallel. To obtain quantitative data, the gel was dried and analyzed using Molecular Imager FX (Bio-Rad). The total extent of RNA cleavage was determined as a ratio of radioactivity found in the tRNA fragments to total radioactivity applied on the gel lane.

To measure the kinetics of the tRNA cleavage with the oligonucleotide conjugates **B-Im(N/m)**, the reaction mixture of 100 μL was prepared containing imidazole buffer, 5×10^{-7} M 3′-[^{32}P]-tRNA^{Phe} and one of the oligonucleotide conjugates **B-Im(N/m)** at different concentrations (from 5×10^{-7} to 5×10^{-4} M). Reaction was performed at 37°C. At indicated time intervals, an aliquot (10 μL) of the reaction mixture was withdrawn and frozen for lately assay. Reaction aliquots were defrosted and immediately precipitated with 150 μL 2% lithium perchlorate in acetone. After that, the reaction mixtures were assayed as described above.

In some experiments, imidazole buffer was replaced by HEPES-buffer (50 mM HEPES-KOH, pH 7.0; 0.2 M KCl, 0.1 mM EDTA, 100 μg/mL RNA-carrier), or Cacodylate buffer (50 mM Cacodylic acid-KOH, pH 7.0; 0.2 M KCl, 0.1 mM EDTA, 100 μg/mL RNA-carrier) or Tris-HCl-buffer (50 mM Tris-HCl, pH 7.0; 0.2 M KCl, 0.1 mM EDTA, 100 μg/mL RNA-carrier). The experiments where imidazole buffer was replaced by other buffers are indicated in text and figure legends.

2.4. Inhibition of Site-Selective tRNA^{Phe} Cleavage by Oligonucleotide Competitor. In the competition experiments, to a reaction mixtures (10 μL) containing imidazole buffer, 5×10^{-7} M 3′-[^{32}P]-tRNA^{Phe} oligonucleotide **A** (complementary to sequence 62-76 of the tRNA^{Phe}) or oligonucleotide **B** was added up to concentration 5×10^{-5} M. The mixtures were incubated for 10 min at 37°C, then one of oligonucleotide conjugates **B-Im(N/m)** at concentration 5×10^{-6} M was added and the mixtures were incubated at 37°C for different times (from 30 min to 2 h). Reaction was quenched and analyzed as described above.

2.5. Gel-Mobility Shift Assay of Hybridization of Oligonucleotide B and Conjugates B-Im(4/2) to tRNA^{Phe}. A gel-mobility shift assay of tRNA^{Phe} binding with oligonucleotide **B** and conjugates **B-Im(4/2)** was performed as follows. [3′-^{32}P] tRNA (5×10^{-7} M) was incubated at 37°C during 30 min in reaction mixture (10 μL) containing hybridization buffer (50 mM Tris-HCl, pH 7.0, 200 mM KCl, 0.1 mM EDTA) and oligonucleotide **B** or conjugate **B-Im(4/2)** ranging in concentration from 1×10^{-7} to 5×10^{-6} M. After incubation, 8 μL of loading buffer (50% glycerol, 0.025% bromophenol blue, 0.025 xylene cyanol) was added to each probe and probes were immediately applied onto running native 10% PAGE with 100 mM Tris-borate, pH 8.3 as running buffer preequilibrated at 4°C for 3 h. Electrophoresis was performed at 4°C for 5 h at 450 V. To obtain quantitative data, the gel was dried and analyzed using Molecular Imager FX (Bio-Rad).

2.6. Molecular Modelling. Molecular modelling was performed using SYBYL 6.6 (TRIPOS Inc.) on a Silicon Graphics O2 workstation as described in [24]. All calculations (MM/MD) were carried out *in vacuo* using Kollman-All force field parameters and a distance-dependent dielectric constant equal to 4. The building of the model molecules (see Figure 1) was carried out in three stages. First, an A-like DNA:RNA hybrid was built using BIOPOLYMER/SYBYL 6.6. Second, the catalytic fragments **Im(4/1)**, **Im(4/2a)**, and **Im(4/2b)** were built using SKETCH/SYBYL 6.6. Charge distributions for these constructs were calculated using a semi-empirical MO program, MOPAC, after geometric optimization. In the last step, each the corresponding cleaving fragment was attached to the DNA:RNA hybrid via the terminal 5′-phosphate group of the G1 nucleotide residue.

For each molecule, several different starting structures were created, differing in the conformation and location of the cleaving group, ranging from "in" to "out" extreme positions relative to the DNA:RNA hybrid. The starting structures of each molecule were subjected to minimization to a gradient convergence of 0.05 Kcal/mol, considering the DNA:RNA hybrid as a solid body. Each starting structure was individually subjected to a 12 ps MD run using simulated annealing as follows. The system was heated at 600 K during 2 ps followed by a gradual cooling down to 100 K over 10 ps. A final minimization to convergence criteria was performed resulting in the respective final conformations for each type of molecule. S and F symbols before the conjugate name indicate the starting and final conformation, respectively.

"Active conformations" were obtained for each model molecule (O1-Im4/1, O2-Im4/1 O1-Im4/2a, O2-Im4/2a, O1-Im4/2b, and O2-Im4/2b, where Type1 shows the anchor group 1) using a similar simulating annealing protocol. In these molecules, the cleaving groups were deliberately located in the most favorable position for successful cleavage at the C_{63}-A_{64} site. Namely, the distance between N1 of one of the imidazole residues and 2′ OH group of C_{63} was ~3 Å and the distance between N1 of the other imidazole residue and

Table 1: Oligonucleotide-based artificial ribonucleases bearing multiple imidazole groups in the catalytic part: structure of conjugates and efficiencies of site-specific cleavage of yeast $tRNA^{Phe}$.

	Conjugate[1]	Number of Im groups	Anchor group[2]	Linker length[3]	Reaction rate, %/1 h[4]	Reaction $_{1/2}\tau$[5]
1	B-Im(2/1)	2	1/2	21	24.8	2 h 15′*
2	B-Im(2/2)	2	2	20	25.3	2 h 5′*
3	B-Im(4/1)	4	1	41	49.8	1 h 5′
4a	B-Im(4/2a)	4	2	29	16.6	2 h 50′
4b	B-Im(4/2b)	4	2	26	5.5	>12 h
5	B-Im(4/3)	4	3	18/22	~1	—[6]
6	B-Im(8/1)	8	1	41	45.1	1 h 15′
7	B-Im(8/2)	8	2	41	29.7	2 h
8	B-Im(8/2+1)	8	1	49	31.3	1 h 45′
9	B-Im(16/1)	16	1	70	38.5	2 h 10′
10	B-Im(8/4+1)	8	4	23/31/41	29.1	2 h
11	B-Im(12/4+2)	12	4	23/31/51	33.8	1 h 40′
12	B-Im(24/4+2)	24	2	41/79	49.9	1 h
13	B-Im(32/4+2)	32	1	41/98	20	2 h 30′

[1]The structures of the conjugates are shown in Figure 1 and in Supplementary material (SM-Figure 1).
[2]The type of anchor groups of the conjugates as shown in Figure 1(a). Type 1: cyclohexanol moiety; 2: 5′-aminothymidine residue; 3: nonnucleotide insertion; 4: 2′-modified uridine residue.
[3]The number of simple C–C, C–N, or P–O bonds between the 5′-terminal phosphate group of oligonucleotide **B** and imidazole groups of RNA-cleaving construct.
[4]Efficiency of the RNA cleavage was measured as percentage of $tRNA^{Phe}$ cleavage at C_{63}-A_{64} phosphodiester bond achieved after incubation during 1 h in the presence of 10 μM of the conjugate.
[5]Time to reach 50% of tRNA cleavage under the standard conditions (see above) at conjugates concentration 10^{-5} M.
[6]50% of $tRNA^{Phe}$ cleavage was not achieved under experimental conditions.
*These conjugates are described in the preceding paper [1].

the oxygen of the phosphate group, connecting C_{63} and A_{64}, was around 3 Å [26, 27]. A final minimization was performed without constraints, using the above protocol.

3. Results and Discussion

3.1. Design of Oligonucleotide-Based Artificial Ribonucleases. The RNA cleaving oligonucleotide conjugates have been prepared in two-step procedure described in [21, 22]. In the first step, the precursor 17-mer GATCGAACACAGGACCT (B) was built up with multiple methoxyoxalamido (MOX) modifiers of different type (see Supplementary Material available online at doi:10.4061/2011/748632 for the structures of all conjugates synthesized in this study). In the second step, the prepared oligonucleotide-MOX precursors were functionalized with histamine to yield after deprotection step the oligonucleotide conjugates **B-Im(N/m)**, bearing from 4 to 32 histamine residues at the 5′-end.

Figure 1 provides structures of oligonucleotide conjugates **B-Im(4/n) and B-Im(8/1)** studied in this work. The conjugates hereafter were referred to as **B-Im(N/m)**, where B-Im designates the oligonucleotide-imidazole conjugates, N shows the number of imidazole residues, and m indicates the type of anchor group. In the cases, when two or three conjugates contain equal number of imidazole residues and the same anchor group, the conjugate number is shown in the brackets. The distance between oligonucleotide and imidazole residues of the cleaving part was varied by altering the structures of the anchor and the linker groups. Cyclohexyl residue was exploited as anchor group in the conjugate Type 1. Thymidine nucleoside was used to anchor the catalytic part in the conjugate Type 2. The thymidine nucleoside, presenting in the structure of conjugate, was supposed to be involved in base pairing with RNA or might be stacked on heteroduplex contributing to the duplex stability. In the conjugate Type 3, an anchor group was chemical construct mimicking phosphodiester bond and used as a rule as nonnucleotide insertion in oligonucleotide. Type 4 anchor group was used in the conjugates **10**–**13** to increase the number of branches, bearing imidazole residues. For this purpose and to increase the linker length and flexibility, the anchor groups of Types 1 and 2 were incorporated also into the linker structures (see Figure 1-SM, conjugates **6**–**9** and **11**–**13**). Moreover, the length and structure of linkers, connecting imidazole residues and anchor, were distinguished sufficiently among the compounds (Table 1 and Figure 1-SM). Conjugates containing equal number of imidazole residues, for example, **B-Im(4/m)** or **B-Im(8/m)**, differ not only in anchor groups, but also in a linker length. The linker of conjugate **B-Im(4/3)** was the shortest one. In this conjugate, four imidazole residues were placed nonsymmetrically in respect to the phosphate groups forming the anchor, which connected the oligonucleotide **B** and the cleaving construct.

The identity of the conjugates was confirmed by mass-spectrometry; (see Figure 1). In addition, the purity of the conjugates was confirmed by electrophoresis in 12% PAAM/8 M urea gel (Figure 1(c)). It is seen that the purity of the conjugates is close to 95%. Only in the case of conjugates **2**,

5, **10**, and **16**, some other bands with the relative intensities of 5–7% are seen.

3.2. Cleavage of Yeast tRNA^{Phe} by the Conjugates Bearing Multiple Imidazole Residues. Yeast tRNA^{Phe} was chosen deliberately as a target to assess cleavage activity of the conjugates. The rationale behind this was that other oligonucleotide-based aRNases including three prepared using precursor approach had been tested using this target [1, 12, 14–16, 28]. Figure 2(a) represents cloverleaf structure of yeast tRNA^{Phe} showing the target site for the conjugates **B-R(N/m)**. The recognition element of the studied aRNases was represented by the 17-mer oligonucleotide **B**, which was complementary to the sequence 44-60 and was designed to deliver the imidazole residues to the sequence ^{61}CACAG^{65} in the T-arm of the tRNA^{Phe}, which is known to be highly sensitive towards cleavage [29–32]. It was previously shown that oligonucleotide **B** could efficiently bind to this sequence in the tRNA^{Phe}, which resulted in unfolding of entire TΨC hairpin, whereas the aminoacceptor and anticodon stems of the tRNA were only slightly affected by the oligonucleotide binding [33–35].

The conjugates **B-Im(N/m)** were designed to catalyze RNA cleavage via transesterification reaction. To compare the ribonuclease activities of the conjugates, we used the conditions of single reaction turnover ([**B-Im(N/m)**]>[RNA]). The oligonucleotide part of the conjugate provides efficient hybridization with target sequence in tRNA^{Phe} and formation of the stable heteroduplex ($K_d = 7\,\mu M \pm 1\,\mu M$; V. Petyuk, unpublished data). We believe that under the experimental conditions the rate and efficiency of site-selective cleavage of the target is affected only by the catalytic part of the conjugate.

Site-selective cleavage of [3′-^{32}P]tRNA^{Phe} by the conjugates was carried out at 37°C in 50 mM imidazole buffer. The reaction was initiated by addition of the conjugate **B-Im(N/m)** to the reaction mixtures. Localization of cleavage sites was performed by comparison of products obtained upon cleavage of tRNA by the conjugates with the products of random RNA cleavage with RNase T1 and 2 M imidazole buffer, pH 7.0 [25, 36]. Figure 2(b) shows typical results of the analysis of tRNA^{Phe} cleavage by the conjugates **B-Im(N/m)** at 37°C under standard conditions. No measurable spontaneous degradation of tRNA under the experimental conditions in the absence of reagents and in the presence of unmodified parent oligonucleotide **B** occurred (Figure 2(a), lanes C1 and C2). The conjugates cleave tRNA^{Phe} with high rate mostly at the phosphodiester bonds C63-A64 and A64-G65. The extents of tRNA^{Phe} cleavage at these bonds are about 70% and 20%, respectively. Moreover, some minor cleavage at the phosphodiester bonds C61-A62 and A62-C63 within the target sequence is observed (Figure 2(a), lanes 1–10). The ratio between the products of tRNA cleavage at these phosphodiester bonds located in close proximity to each other in the target sequence seems to correspond to steric accessibility of these bonds to cleavage by the conjugates B-Im(N/m). In the case of conjugate **B-Im(4/2a)** (similar cleavages were observed for the conjugates with 16, 24, and 32 imidazole residues), slight cleavages are observed at U8-A9 and C75-A76 phosphodiester bonds (less then 3% of the total extent of tRNA cleavage), known as well as C61-A62 and C63-A64 bonds in the yeast tRNA^{Phe} to be very sensitive toward cleavage by various agents [29–32]. The bond U8-A9 is located in the proximity to sequence C61-G65 if tertiary structure of tRNA^{Phe} is taken into account. Thus, one may assume that the observed cleavage of this bond occurs in complementary complex formed by conjugate **B-Im(N/m)** with tRNA and is the result of flexibility of RNA-cleaving group of the conjugates. Besides, we can assume that initial cleavage of phosphodiester bond C63-A64 produces a nick in tRNA molecule that may increase the flexibility of the complex tRNA-conjugate **B-Im(N/m)** and thus makes phosphodiester bonds C64-G65 and U8-A9 available for catalyst's function groups. The slight cleavage observed at C75-A76 site apparently is the result of spontaneous tRNA hydrolysis in the duplex with the conjugate.

Competition experiments were used to prove that tRNA cleavage occurs in *tRNA-conjugate* complex. tRNA^{Phe} cleavage by conjugate **B-Im(4/1)** was performed in the presence of oligonucleotide inhibitors of two types: parent oligonucleotide **B** and oligonucleotide **A** complementary to the sequence 61-75 (Figure 2(b), lanes 17–19). As expected, the cleavage of tRNA^{Phe} by conjugate **B-Im(4/1)** is considerably decreased in the presence of oligonucleotides **B** (primary data not shown) and is completely abolished in the presence of oligonucleotide **A**, which protects target sequence by duplex formation. It is worth noting that in the presence of oligonucleotides **B**, the cleavage of tRNA^{Phe} is suppressed in all sites due to competitive binding, whereas oligonucleotides **A** inhibit tRNA^{Phe} cleavage only at the target sequence failing to change reaction rate at other sites. These results indicate the structural specificity of cleavage and stress the fact that the cleavage occurs only at a single-stranded target sequence.

Figure 2(c) shows concentration dependencies of the cleavage reaction for conjugates **B-Im(4/1)** (**3**) and **Im(24/4+2)** (**11**) (curves 1 and 2, resp.) and binding of the conjugate **B-Im(4/2a)** and parent oligonucleotide **B** with tRNA (curves 3 and 4, resp.). It is seen that the cleavage data are in agreement with the complex formation. Furthermore, the binding affinity of the parent oligonucleotide and the conjugate studied under identical conditions are found to be similar (Figure 2(c)). Therefore, the conjugation of bulky imidazole-containing reactive groups to the oligonucleotide does not affect the hybridization process.

To estimate the influence of buffer on efficiency of site-selective RNA cleavage, we compared tRNA^{Phe} cleavage by the conjugate **B-Im(4/2)** in different buffer solutions (Figure 2(d)), taken 50 mM imidazole buffer, pH 7.0 as a standard conditions. The rate of site-selective cleavage is affected by replacement of 50 mM imidazole buffer, pH 7.0, by 50 mM Tris-HCl buffer or 50 mM cacodylate buffer at pH 7.0 by the factor of 1.5 indicating that imidazole itself may contribute to cleavage reaction. At longer incubation time, this difference became insignificant and does not exceed estimated experimental error. Surprisingly, 50 mM HEPES buffer entirely abolishes cleavage inactivating the conjugates. The reason of this is not entirely understood.

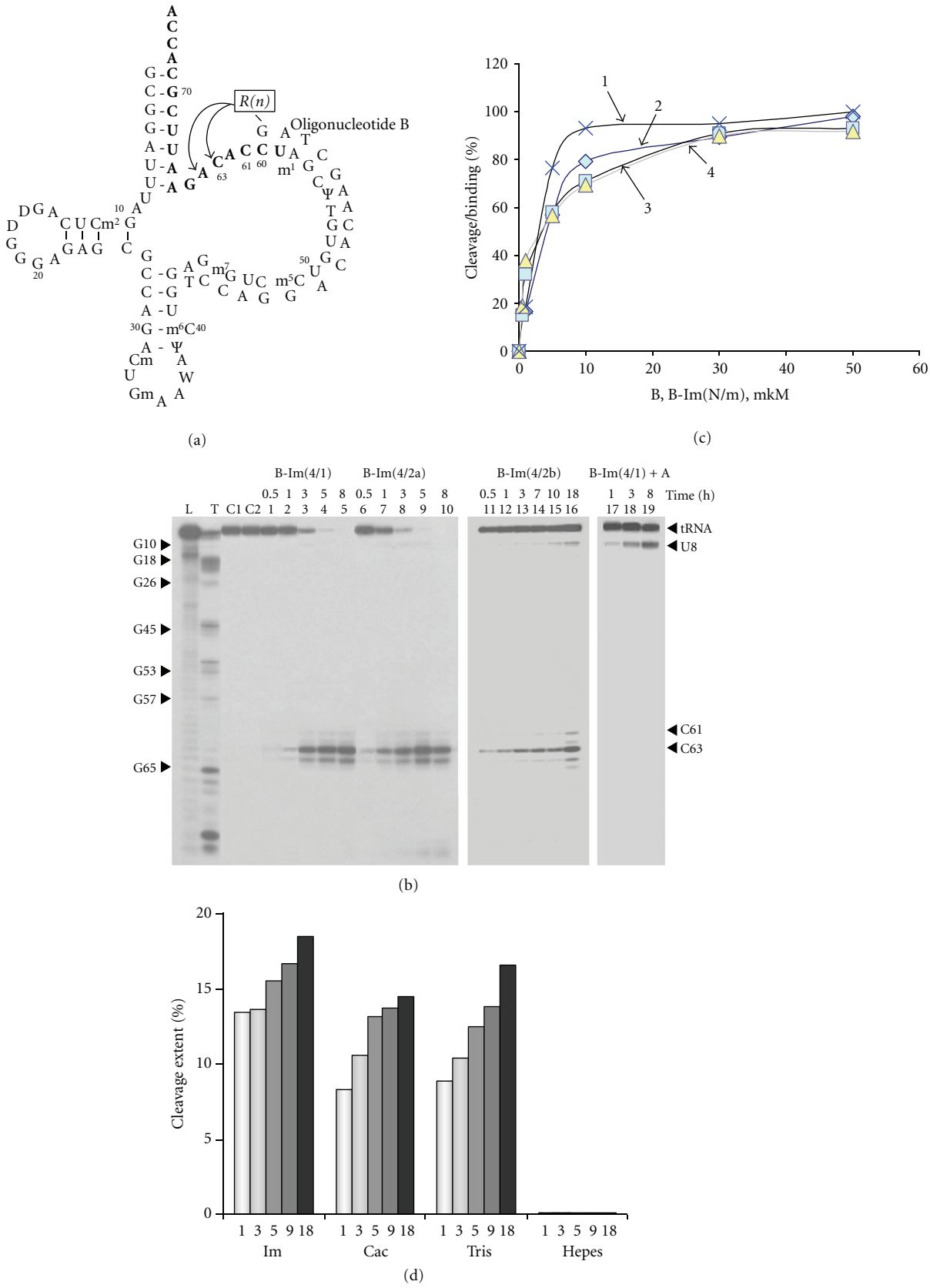


Figure 2: *Site-selective cleavage of yeast tRNA^{Phe} by the oligonucleotide conjugates bearing multiple imidazole residues in the catalytic part.*

One explanation could be that sulfonic acid of HEPES [4-(2-hydroxyethyl)-piperazine-1-ethanesulfonic acid] may interact with protonated imidazole residues of the conjugates and inactivate it. Previously, we also observed twofold decrease of the rate of RNA cleavage by oligonucleotide conjugate bearing two imidazole residues in 50 mM HEPES-KOH buffer, whereas the other buffers only slightly affected the rate of RNA cleavage [37].

In Figure 2, (a) represents cloverleaf structure of yeast $tRNA^{Phe}$ and target sequences for oligonucleotide **B**-based artificial ribonucleases **B-Im(N/m)**. Bold letters show target sequence for oligonucleotide-competitor **A**. Arrows indicate the sites of site-selective tRNA cleavage by the conjugates; (b) represents cleavage of $tRNA^{Phe}$ with oligonucleotide conjugates **B-Im(4/m)**. Autoradiograph of 12% polyacrylamide /8 M urea gel. Cleavage reactions were performed at 37°C in 50 mM imidazole buffer, pH 7.0, containing 200 mM KCl, 0.2 mM EDTA, 100 μg/mL RNA carrier, 5×10^{-7} M [3′-^{32}P]$tRNA^{Phe}$ and 1×10^{-5} M one of the conjugates **B-Im(4/1)**, **B-Im(4/2a)**, and **B-Im(4/2b)**. Lanes L and T1 are Imidazole and RNase T1 ladders, respectively. Lanes C1 and C2 are tRNA incubated in the reaction buffer for 8 h in the absence and in the presence of oligonucleotide **B** (5×10^{-5} M), respectively. tRNA cleavage by conjugate **B-Im(4/1)** (**3**) for 0.5, 1, 3, 5, and 8 h (lanes 1–5, resp.), by conjugate **B-Im(4/2a)** (**4a**) for 0.5, 1, 3, 5, and 8 h (lanes 6–10), by the conjugate **B-Im(4/2b)** (**4b**) for 0.5, 1, 3, 7, 10, and 18 h (lanes 11–16); by the conjugate **B-Im(4/1)** (**3**) in the presence of oligonucleotide **A** (1×10^{-4}) for 0.5, 1, and 5 h (lanes 17–19). Concentration of the conjugates in the reaction mixtures was 1×10^{-5} M. Reactions were quenched and analyzed as described in the experimental part; **(c)** represents concentration dependencies of the $tRNA^{Phe}$ cleavage by the conjugates **B-Im(4/1)** (**3**) and **B-Im(24/2)** (**12**) (curves 1 and 2, resp.) and complex formation with the conjugate **B-Im(4/2a)** (**4a**) (curve 3) and parent oligonucleotide **B** (curve 4). The reaction mixtures were incubated for 2 h under the conditions described above; **(d)** represents the influence of the buffer nature on the rate of $tRNA^{Phe}$ cleavage by the conjugate **B-Im(4/2a)** (**4a**). [3′-^{32}P]$tRNA^{Phe}$ (0.5 μM) was incubated with 2 μM of (**4a**), in 50 mM imidazole buffer, pH 7.0 (Im) or 50 mM cacodylate buffer, pH 7.0 (Cac) or 50 mM Tris-HCl buffer, pH 7.0 (Tris) or 50 mM HEPES buffer, pH 7.0 (Hepes) containing 0.2 M KCl, 0.2 mM EDTA, and 0.1 mg/mL RNA carrier.

3.3. Kinetics of $tRNA^{Phe}$ Cleavage by the Conjugates. Figure 3(a) displays the kinetics of $tRNA^{Phe}$ cleavage with conjugates under the study taken at conjugate concentration 10 μM. It is seen that the conjugates cleave RNA with different efficacy. One of the most effective conjugate is **B-Im(4/1)**, as complete site-selective cleavage of the target by this conjugate is achieved within 2-3 hours. Similar kinetics are displayed by the conjugate **B-Im(8/1)** (**6**) and the **B-Im(24/4+2)** (**12**). The conjugates **B-Im(4/2a)** (**4a**) and **B-Im(8/2)** (**7**) catalyze RNA cleavage less efficiently as compared with **B-Im(4/1)** or **B-Im(8/1)**. Cleavage efficiency of the conjugate **B-Im(4/3)** (**5**) was essentially poor.

Figure 3(b) shows kinetic curves of $tRNA^{Phe}$ cleavage by conjugate **B-Im(4/1)** obtained at different concentrations of the conjugate. The character of kinetic curves evidences the bimolecular reaction: the rate of RNA cleavage is increased with the increasing of the conjugate concentrations. At concentration of conjugate **B-Im(4/1)** of 10 μM, which provides its almost quantitative binding to tRNA, the half-life of tRNA is about 1 h (Figure 3(b), curve "10 μM"). When concentration of conjugate **B-Im(4/1)** is increased up to 50 μM, the half-life of tRNA is counted by minutes (Figure 3(b), curve "50 μM"). It should be noted that the site-selective cleavage of tRNA proceeds much slower than the binding of the parent oligonucleotide **B** to tRNA under similar conditions [33–35]. At concentration of the conjugate equal to or twice as higher as concentration of the tRNA, the curves reveal lag period (Figure 3(b), curves "0.5" and "1 μM," resp.) indicating that the process of site-selective RNA cleavage includes several consecutive steps occurring with comparable rates. Another explanation of this could be that the reaction has more complex character and goes through some stages other than binding of the conjugate and cleavage, which results in sigmoid-like kinetic curve.

For conjugate **B-Im(4/1)**, we calculated the association rate and cleavage rate constants from the data shown in Figure 3(b), assuming that cleavage of the tRNA occurs within the complex with the conjugate. In this case, hybridization step is the second-order reaction and cleavage stage is intramolecular first-order reaction. It turns out that cleavage *pe se* occurs 50 times faster than the binding of the conjugate with the complementary sequence within RNA target (association rate constant and cleavage rate constant are (0.074 ± 0.002) $M^{-1}\,s^{-1}$ and $(3.4 \pm 0.55) \times 10^{-3}\,s^{-1}$, resp.). In other word, binding of the conjugate with tRNA is the rate-limiting stage of site-selective RNA cleavage.

3.4. Correlations between Structure and Activity of the Conjugates Bearing Multiple Imidazole Groups. The conjugates display different cleavage activities (Table 1) ranging from very high cleavage rate ($\tau_{1/2} \sim 1$ h in the case of **B-Im(4/1)**) to very low ($\tau_{1/2}$ value was not achieved under the experimental conditions in the case of **B-Im(4/3)**). Remarkable cleaving activity of the conjugates **3**, **6**, and **12** as well as moderate cleavage activity of the conjugates **4a**, **8**, **10**, **11**, and **13** and poor efficiency of the conjugate **5** are likely resulted from the differences in the structure of the catalytic part of the conjugates rather than due to different hybridization properties and the cleavage reaction *pe se*. The conjugates under the study contain identical oligonucleotide recognition element and display similar cleavage specificity with only minor differences in the cleavage pattern. Therefore, neither hybridization nor sensitivity of the phosphodiester bonds toward the cleavage could explain the differences in the catalytic properties of the conjugates. Even the catalytic parts of the conjugates in several cases contain the equal number of imidazole residues.

Table 1 summarizes data on the cleavage activity and structure of the conjugates under the study (linker length in

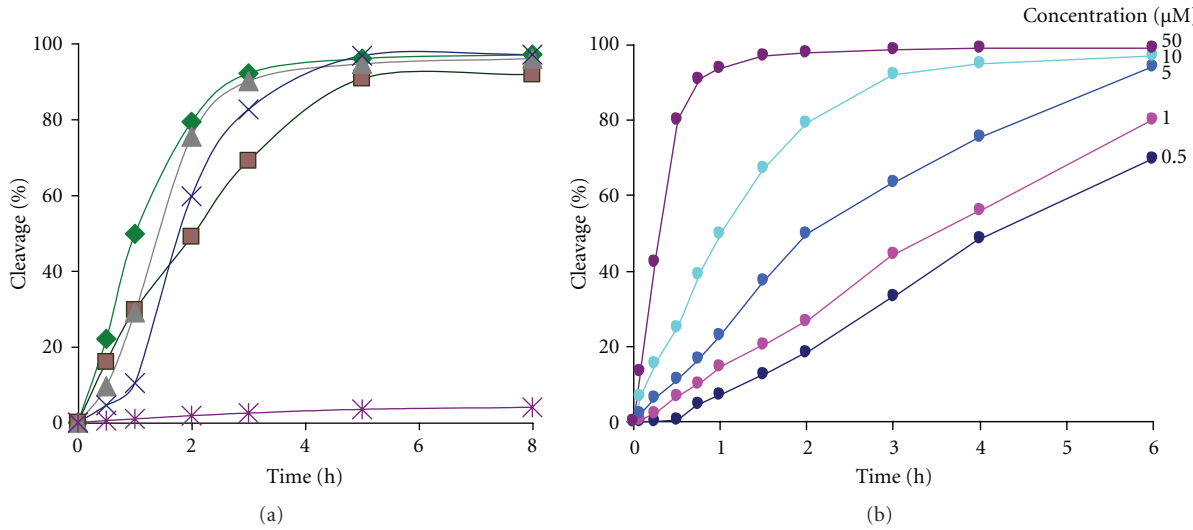


FIGURE 3: *Kinetics of tRNA* ^{Phe} *cleavage by the conjugates:* (a) tRNA cleavage by **B-Im(4/1)** (**3**) (rhombs), **B-Im(4/2a)** (**4a**) (squares), **B-Im(4/3)** (**5**), (asterisks), **B-Im(8/4+1)** (**10**) (triangles), and **B-Im(24/4+2)** (**12**) (crosses) under standard conditions at conjugates concentration 10 μM. (b) tRNA cleavage by **B-Im(4/1)** (**3**) at concentrations ranging from 0.5 to 50 μM.

total number of simple C–C, C–N, C–O, and P–O bonds between 5′-end of oligonucleotide B and imidazole residues, type of anchor group, and reaction halftimes). It is seen that hydrolytic activity of compounds is in a good agreement with the length of their linkers (Table 1). Amongst the conjugates bearing four imidazole residues, the highest cleavage activity is observed for the conjugate **B-Im(4/1)**, which has the linkers of 41 simple bonds. The conjugates **B-Im(4/2a)** with the linkers of 29 chemical bonds exhibit 3-fold lower cleavage activity, whereas **B-Im(4/3)** (18/24 bonds) cleaves tRNA with lowest efficiency. Thus, stepwise shortening of a linker part of conjugate **B-Im(4/m)** from 41 to 18/24 simple bonds results in decreasing of cleavage efficiency from 50% per hour to 1% per hour. As it is seen from Table 1, only the conjugates containing the linkers of 40 simple bonds are capable of efficiently catalyzing cleavage of $tRNA^{Phe}$ in a site-selective manner. Thus, we can conclude that a long flexible linker of 40 or more simple C–C, C–N, or P–O bonds is required for efficient site-selective cleavage of RNA target.

Figure 4 display *Structure-Activity* correlations found for the conjugates **B-Im(N/m)**. It is seen that cleavage activity of the conjugates **B-Im(N/1)** depends on the number of imidazole groups in the catalytic part (Figure 4(a)). The conjugates bearing 4 and 8 imidazole residues exhibit the highest cleavage activity, while increasing of the number of imidazole groups inhibits cleavage. It is worth pointing out that sigmoidicity of kinetic curves taken under identical conditions is increased with increasing of the number of imidazoles in the conjugate (primary data not shown). It is likely that cleavage activity of these conjugates is affected by a steric interference between imidazoles, thus leading to a loss of cleavage rate. Therefore, four (and/or eight) imidazole residues in the catalytic part of the conjugate seem to be sufficient for efficient RNA cleavage.

From the data shown in Figure 4(b), it is seen that ribonuclease activity of the conjugates depends on type of the anchor group used for attachment of parent oligonucleotide to RNA-cleaving construct. The conjugates with Type 1 anchor group (**B-Im(N/1)**), cyclohexyl moiety (Figure 1(a)), display higher cleavage activity than conjugates with the anchor group in a form of thymidine residue (Type 2, Figure 1) and especially than conjugates with a nonnucleoside anchor group (Type 3, Figure 1) (Table 1). The difference between cleavage activity of the conjugates with Type 2 and Type 4 anchor groups does not exceed experimental error, but is reliably lower than that for conjugates with Type 1 anchor group. Thus, the structure and flexibility of the linker part, the number of imidazole residues in the catalytic part, and the type of anchor group affect cleavage activity of the conjugates and play a role in the optimal positioning of the catalyst in respect to scissile phosphodiester bond.

3.5. Molecular Modeling Study. The main goal of the molecular modelling was to clarify the following issues.

(1) Is there any preferable orientation(s) of the cleaving group(s) in the vicinity of the cleavage site, which could explain the hydrolytic activity (or its absence) of each conjugate?

(2) Is there any evidence of interaction(s) between the cleaving constructions and the DNA:RNA hybrid, which could stabilize preferable orientation(s) of these groups near the cleavage site?

(3) What is the correlation between the structure of the cleaving constructs and the linker groups and their hydrolytic activity?

It is evident from the biochemical assays that hydrolytic activity within the heteroduplex *tRNA* ^{Phe}*-conjugate* is governed by the intrinsic properties of the cleaving constructs, but not the detailed structure of the DNA-RNA hybrid, which was the same in all experiments. Therefore, the main priority was to identify possible location(s) of the cleaving

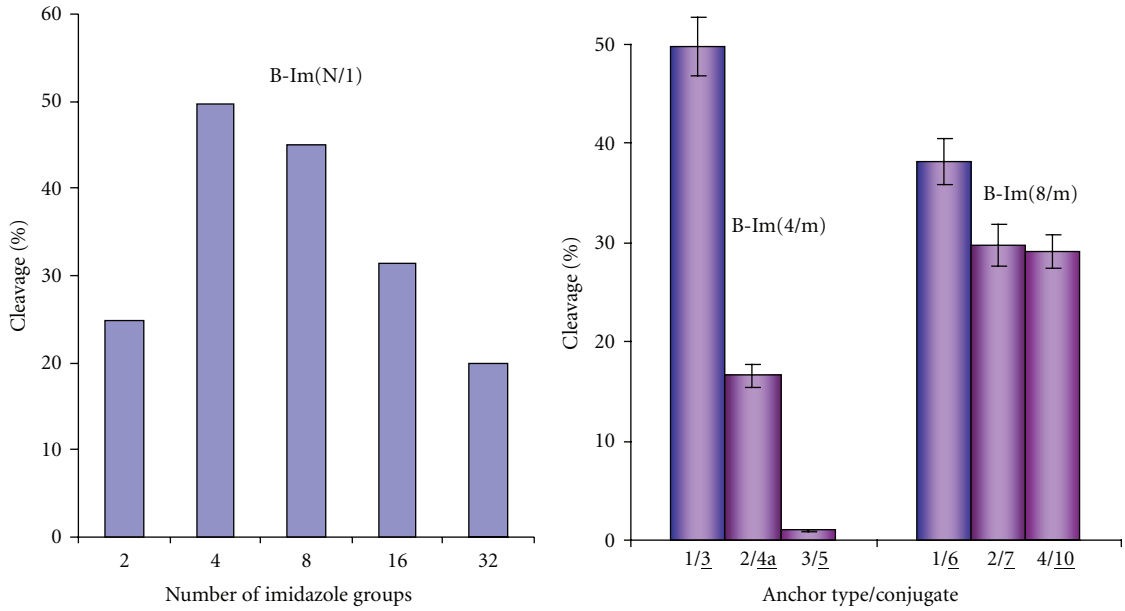


Figure 4: *Structure-activity correlations found for the conjugates* ***B-Im(N/m)***. (a) Correlations between the number of imidazole groups in catalytic parts of conjugates and their catalytic activities. Cleaving data corresponds to incubation of the RNA with conjugates **1**, **3**, **6**, **9**, and **13** under standard conditions for 1 h. (b) Correlation between the type of anchor groups and catalytic activity of conjugates. Diagram presents data for tetraimidazole-containing conjugates **3**, **4a**, **5** (**B-Im(4/m)**) and octaimidazoles-containing conjugates **6**, **7**, **10** (**B-Im(8/m)**).

group(s) relatively to the cleavage site rather than to determine the detailed conformation of the hybrid (which would be impossible in the absence of relevant NMR/X-ray data).

It has been shown earlier that hybridization of long oligonucleotides to TΨC-loop [33, 38] or to the 3′-acceptor stem [33–35] of tRNA induced a local unfolding of the tRNA secondary structure. Therefore, we assumed that the hybrid part of the molecules under the study can be represented to a first approximation by a double-stranded DNA:RNA duplex, which would form primarily an A-like conformation, according to physicochemical investigations [39–41]. In our molecular modeling studies, we used a shortened model for the heteroduplex, comprising the target RNA sequence (corresponding to the $^{5'}$A44-U69 fragment of tRNA^{Phe}) and the oligodeoxyribonucleotide conjugate complementary to this RNA region (Figure 2(a)).

Based on this hybrid duplex, three model molecules were created (M-Im(4/1), M-Im(4/2a), and M-Im(4/2b)) by attachment of the respective cleaving constructs to the 5′ terminal phosphate group of the G^1 nucleotide residue (see Figures 1 and 2). At this stage, in the absence of the detailed structural data for the hybridized heteroduplex, it was reasonable to treat the hybrid as a "solid" structure [24] whilst allowing cleaving constructs to possess a full conformational flexibility during the calculations.

For each molecule M-Im(4/1), M-Im(4/2a), and M-Im(4/2b), several starting structures were created differing in the spatial orientation of the cleaving construct(s), ranging from "in" to "out" extremes (see "Section 2"). This increased the conformational space searched for each model molecule. All starting structures were subjected to independent simulating annealing (see "Section 2") giving rise to the respective final structures (M-Im(4/1): F1-Im(4/1)–F9-Im(4/1); M-Im(4/2a): F1-Im(4/2a)–F9-Im(4/2a); M-Im(4/2b): F1-Im(4/2b)–F9-Im(4/2b).

Both structural parameters and energy values of each final structure were analyzed and compared with those of "active" conformations (O1-Im(4/1), O2-Im(4/1), O1-Im(4/2a), O2-Im(4/2a), O1-Im(4/2b), and O2-Im(4/2b)), which had deliberately preorganized cleaving centers according to [20, 26, 27, 42, 43] (see "Section 2") (Table 2).

Molecules containing four imidazole residues M-Im(4/1), M-Im(4/2a), and M-Im(4/2b), having two bis-imidazole cleaving constructs, differ in the geometrical properties of their linker groups (see Figure 1). The linker group of M-Im(4/1) possess a cyclohexyl fragment connecting the cleaving moieties with the oligonucleotide. In M-Im(4/2a), the cyclohexyl fragment is replaced by a deoxyribothymidine fragment. M-Im(4/2b) is a shorter analogue of M-Im(4/2a).

For each molecule, nine final structures plus two "optimal" conformations were obtained as a result of the simulated annealing calculations. The data on the distances between N1 atoms of imidazole rings and target site as well as energy values are presented in Table 2.

For M-Im(4/1) (Table 2), the most energetically favorable conformations (O1-Im(4/1), O2-Im(4/1), and F4-Im(4/1)) are characterized by the location of at least two imidazole groups close to the target site (3–7 Å). In the most cases, M-Im(4/2a) conformers have at least one imidazole group located near the cleavage site, which suggests a high probability for cleaving groups to form a preorganized

Table 2: Distances between active centers of the conjugates **B-Im(4/m)** and RNA target atoms.

Molecule	Final structure	Energy (kcal/mol)	ΔE (kcal/mol)		Dist. Pair 1(E)		Dist. Pair 2(E)	
			for O1	for O2	Im1-OH[(a)]	Im2-OP[(b)]	Im1-OH	Im2-OP
M-Im(4/1)	F1- Im(4/1)	47.3	18.2	13	6.2	11.4	26.6	21
	F2- Im(4/1)	66.7	37.6	32.4	6.2	14.3	21.1	22.9
	F3- Im(4/1)	48.3	19.2	14.0	6.6	16.3	23.0	33.9
	F4- Im(4/1)	28.6	−0.6	−5.8	7.1	7.0	2.9	7.5
	F5- Im(4/1)	60.2	31.1	25.9	7.5	17.2	11.8	11.5
	F6- Im(4/1)	42.6	13.5	8.3	12.8	9	23.1	13.4
	F7- Im(4/1)	54.9	25.8	20.5	11.9	12.7	15.7	25.9
	F8- Im(4/1)	47.3	18.2	13.0	7.3	15.0	24.8	21.0
	F9- Im(4/1)	49.8	20.7	15.5	10.7	17.0	16.5	23.7
	O1- Im(4/1)	**29.1**	—	—	**3.0**	**3.0**	12.9	23.1
	O2- Im(4/1)	**34.3**	—	—	23.5	25.7	**3.2**	**3.3**
M-Im(4/2a)	F1- Im(4/2a)	−6.7	−4.3	−5.8	12.4	25.1	26.4	19.8
	F2- Im(4/2a)	24.0	26.3	24.8	23.1	15.6	23.2	34.8
	F3- Im(4/2a)	0.5	2.9	1.4	21.1	11.6	21.3	19.8
	F4- Im(4/2a)	9.7	12.0	10.5	6.4	15.8	23.3	27.8
	F5- Im(4/2a)	1.7	4.0	2.5	5.6	10.1	21.6	26.7
	F6- Im(4/2a)	23.4	25.8	24.3	26.1	30.6	16.2	8.6
	F7- Im(4/2a)	23.4	25.8	24.3	26.1	30.6	16.2	8.6
	F8- Im(4/2a)	0.5	2.9	1.4	21.2	11.6	24.4	17.4
	F9- Im(4/2a)	13.3	15.6	14.1	14.4	18.4	28.2	27.8
	O1- Im(4/2a)	**−2.4**	—	—	**3.17**	**3.0**	20.1	33.7
	O2- Im(4/2a)	**−0.9**	—	—	17.7	17.6	**3.0**	**3.2**
M-Im(4/2b)	F1- Im(4/2b)	−4.7	6.9	−16.2	23.7	15.9	21.4	19.9
	F2- Im(4/2b)	−17.9	−6.3	−29.3	30.9	32.7	21.9	30.7
	F3- Im(4/2b)	−17.8	−6.2	−29.2	27.6	36.0	25.9	21.9
	F4- Im(4/2b)	0.1	11.7	−11.3	5.3	23.3	14.5	14.3
	F5- Im(4/2b)	−9.1	2.5	−20.5	21.1	26.2	23.5	25.8
	F6- Im(4/2b)	0.1	11.7	−11.3	18.5	9.2	11.3	19.7
	F7- Im(4/2b)	0.1	11.7	−11.3	18.5	9.2	11.3	19.7
	F8- Im(4/2b)	−6.7	4.9	−18.1	9.0	10.6	20.7	23.3
	F9- Im(4/2b)	−2.5	9.1	−13.9	14.9	30.4	27.9	31.8
	O1-Type3	**−11.6**	—	—	**3.1**	**3.0**	13.4	14.9
	O2-Type3	**11.4**	—	—	25.2	27.2	**3.0**	**3.0**

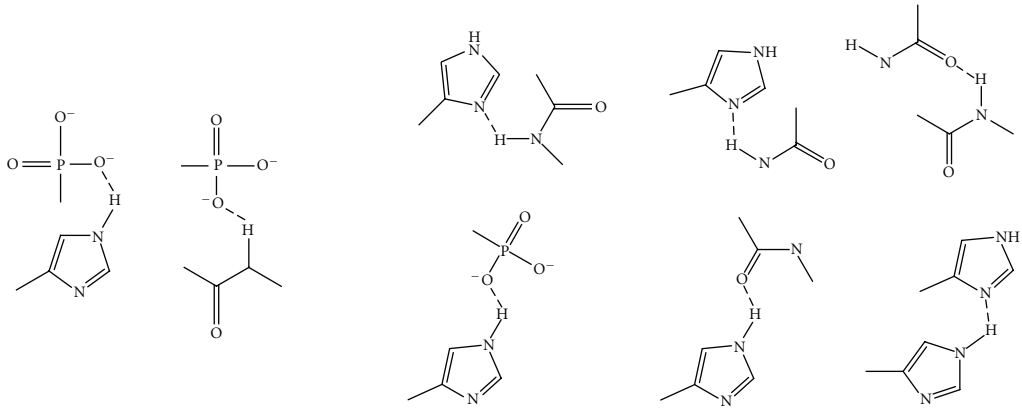
[(a)] Im1-OH represents the distance between N1 of one of the imidazole residues and the 2′OH group of C^{63}.
[(b)] Im2-OP represents the distance between N1 of the other imidazole residue and the oxygen of the phosphate group connecting C^{63} and A^{64}.

"active" conformation. These data are in agreement with the biochemical assays on high hydrolytic activity seen for **B-Im(4/2a)** conjugate, showing that this compound reaches the maximum activity (49.8%) found for artificial ribonucleases. Figure 5 represents the fragments of the F4-Im(4/1) structure, as the most energetically favorable conformation of M-Im(4/1), and shows the probable orientation of the cleaving groups.

In the case of M-Im(4/2a), most of the final structures obtained in the simulated annealing experiments are characterized by a distant location of cleaving groups from the target site (Table 2). For example, the lowest-energy structures F1-Im(4/2a), F3-Im(4/2a), and F8-Im(4/2a) have very unfavorable orientations of imidazole cleaving groups in terms of hydrolytic activity (see Figure 6 showing F1-Im(4/2a) final structure as an example). The possible reason of this could be the presence of a deoxyribothymidine fragment linking the oligonucleotide and the cleaving construct. According to our data, the thymidine base is involved in the DNA:RNA helix structure in a semistacking interaction with the neighboring nucleotide residue dG^1. The thymidine base matches comfortably the duplex structure, although no additional hydrogen bonding with the opposite strand was detected (for instance, with the A^{62}). This interaction could control the overall conformation of the rest of the cleaving construct and possibly restrict the conformational freedom of this group, resulting in a decreased probability to render an "active" conformation. This could explain the relatively

Type 1: intermolecular hydrogen bonds.

Type 2: intramolecular hydrogen bonds.



(a) (b)

Scheme 1

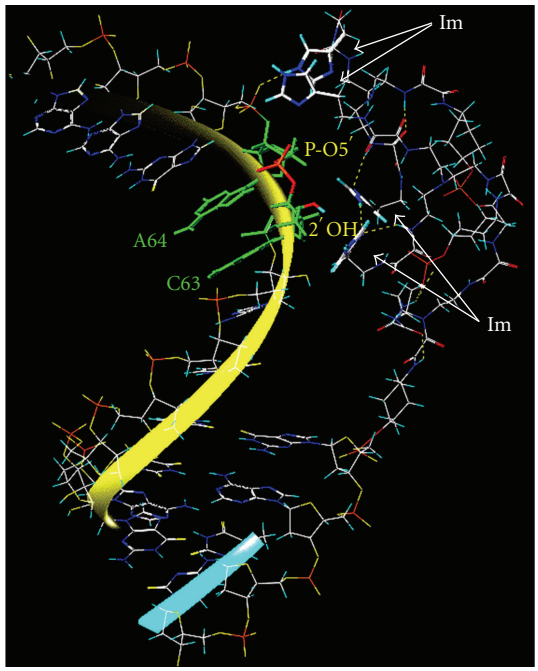


Figure 5: A fragment of the most favourable conformation, F4-Im(4/1), found for hybrid complex H- Im(4/1) between artificial ribonuclease **B-Im(4/1)** and $tRNA^{Phe}$ where the close proximity of imidazole cleaving groups to the target sequence C^{63}-A^{64} (green) is evident. Hydrogen bonds are shown in yellow.

low hydrolytic activity of **B-Im(4/2a)** (16.6%) compared to **B-Im(4/1)**.

Even more pronounced decrease in hydrolytic activity (down to 5–7%) was observed for conjugate **B-Im(4/2b)**, a structural analogue of **B-Im(4/2a)** with a shortened linker group. Data obtained in biochemical assays are consistent with results from molecular modeling for M-Im(4/2b) (Table 2), showing that the most energetically favorable conformations (for instance, F2-Im(4/2b), F3-Im(4/2b), and F5-Im(4/2b)) have a distant location of the cleaving groups relative to the target site (Figure 7 shows F5-Im(4/2b)). Moreover, only one of two bis-imidazole groups can occupy a position, favorable for cleaving activity. This decreases the statistical probability of reaching an "active" conformation for M-Im(4/2b). Generally, most of the final structures of M-Im(4/2b) are characterized by long distances between imidazole rings and target atoms. Analysis of the structural parameters showed that the length of the linker group is not enough to provide a good chance to reach the target site. This possibly contributes to the restricted conformational freedom caused by thymidine interactions with the RNA:DNA duplex structure (see above for M-Im(4/2a)). Taken together, these explain the lowest hydrolytic activity for **B-Im(4/2b)** between the artificial ribonucleases.

Analysis of final structures resulting from the simulating annealing experiments shows a remarkable ability of all cleaving constructs to form multiple hydrogen bonds that generally can be represented by two types (Scheme 1). The first type is the intermolecular interactions between cleaving constructs and the RNA target, which are formed by oxygen atoms of the RNA phosphate backbone and either (i) by imidazole rings of the cleaving groups or (ii) by protons of the linker amido groups (see Scheme 1). These types of hydrogen bonding seem to be in favor of the hydrolytic activity of compounds due to their ability to stabilize the location of cleaving groups, close to the RNA target. The second type of possible hydrogen bonding (Scheme 1) refers to intramolecular interactions within cleaving constructs themselves and involves interactions between imidazole rings and amido groups, carbonyl groups and phosphate groups of the linker. This type of hydrogen bond interaction also includes intralinker bridges formed by amido and carbonyl/phosphate groups. Intramolecular hydrogen bonds seem to compete with favorable intermolecular interactions, preventing the imidazole groups from appropriate contacts with

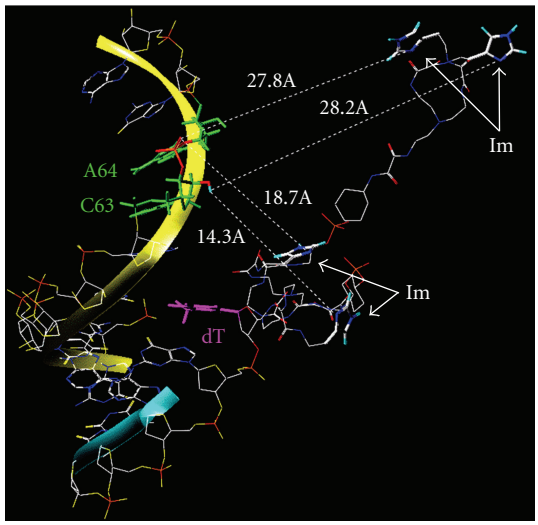


Figure 6: A fragment showing one possible conformation (namely, final structure F9-Im(4/2a)) of **B-Im(4/2a)** (**4a**) found for the hybrid complex H-Im(4/2a). Distances between the cleaving groups and the sequence C^{63}-A^{64} (green) are indicated. The dT (magenta) semistacking interaction with the dG^{1}fragment is also shown. Hydrogen atoms have been removed for easier visualization.

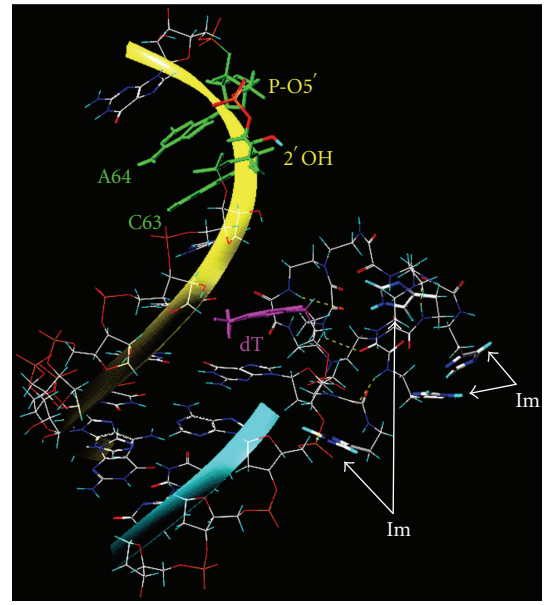


Figure 7: A fragment showing one possible conformation for the hybrid complex between artificial ribonuclease B-Im(4/2b) and tRNA. In this final conformation, F5-Im(4/2b), the imidazole cleaving groups are located distant from the target sequence. The dT linker fragment (magenta) shows interaction with the dG^{1} residue in a semi-stacking interaction, possibly restricting the conformational freedom of this molecule. Hydrogen bonds are shown in yellow.

the RNA backbone, which may result in a decrease of their hydrolytic activity.

Based on these results, we propose several suggestions on how to improve further the cleaving potential of artificial ribonucleases. First, it is important to avoid the use of any carbonyl/phosphate groups within the linker in order to decrease the possibility of intramolecular hydrogen bonding. Secondly, it could be advantageous to introduce polycationic groups within the linker, which might increase possible contacts of cleaving groups with negatively charged phosphate backbones of target RNA. Finally, the insertion of a suitable hydrophobic/intercalating groups within the linker might also increase chances for these groups to be located in the vicinity of the RNA target due to additional stacking/hydrophobic interactions with different ribonucleotide environments.

4. Conclusions

In the current research, we studied oligonucleotide-based artificial ribonucleases containing *multiple* imidazole residues in the catalytic part of the aRNases with systematically varied structure of cleaving constructs. All the conjugates contain the same addressing oligonucleotide which provides efficient and almost quantitative binding of the conjugates to the target sequence within the tRNA. This allows comparing different cleaving constructs in terms of site specificity and efficiency of RNA cleavage. Obtained results let us conclude that the efficiency of site-selective RNA cleavage is governed by a number of dynamic parameters among which the most important is the flexibility of a linker between the catalytic imidazole residues and the addressing oligonucleotide. The other factors affecting the efficiency of site-selective RNA cleavage are the number of imidazole residues in the catalytic part, the type of anchor group, connecting linker structure and the oligonucleotide, and the length of a linker between the catalytic imidazole groups of the construct and the oligonucleotide. We found that four imidazole residues located in the catalytic part of the conjugate can efficiently catalysed cleavage of phosphodiester bond in the case when they are located in the close proximity of RNA sugar-phosphate backbone, while increasing of the number of imidazole groups inhibits cleavage, possibly due to steric interference between the imidazole residues.

Molecular modeling carried out to explain differences in ribonuclease activity of the conjugates showed that preferable orientation(s) of cleaving constructs strongly depend on the structure of the anchor group and length of the linker. The inclusion of deoxyribothymidine as an anchor group significantly reduced the probability of cleaving groups to locate near the cleavage site due to a stacking interaction with the neighbouring nucleotide residue. Remarkably high cleavage activity was displayed by the conjugates with the most flexible and extended cleaving construct, which presumably provides a better opportunity for imidazole residues to be correctly positioned in the vicinity of scissile phosphodiester bond.

The other very important factor affecting the efficiency of site-selective RNA cleavage is sensitivity of phosphodiester bonds of RNA toward cleavage. Different catalytic structures display some sequence preference [1]. On the other hand, phosphodiester bonds in different RNA sequence display different sensitivities to cleaving agents [3, 44–48]. These factors should be taken into account when selecting target

sites within specific RNAs. In our experiments, the cleaving groups were targeted to CACA sequence of $tRNA^{Phe}$, known to be of particular sensitivity to cleavage by different agents. So, the ribonuclease activity of the conjugates was compared under favourable conditions. On the other hand, binding of oligonucleotide conjugates to target RNA is followed by RNA structure rearrangements that can bring the target site in favorable or unfavorable positions toward the catalytic groups [32, 49, 50]. Thus, optimisation of the structure of catalytic groups of oligonucleotide conjugates and identification of optimal structures of RNA for targeting with artificial ribonucleases may facilitate the creation of highly specific artificial ribonucleases of a new generation.

Results of the studies performed by the authors revealed the problems to be solved in order to develop the highly efficient RNA cleaving oligonucleotide conjugates. A few important factors determining efficacy of RNA cleavage remain to be investigated so that the oligonucleotide conjugates take advantage of the known mechanisms used by natural enzymes for achieving high reaction rate.

Acknowledgments

This work was supported by the grants from The Welcome Trust (no. 063630), RFBR (nos. 02-04-48664 and 03-04-06238), RAS program "Molecular and cellular biology," Ministry of Science and Education of the Russian Federation (state contract P438), and President's Program in support of leading scientific schools SS-7101.2010.4.

References

[1] M. A. Zenkova and N. G. Beloglazova, "Site-specific artificial ribonucleases: conjugates of oligonucleotides with catalytic groups," in *Nucleic Acids and Molecular Biology: Artificial Nucleases*, M. A. Zenkova, Ed., vol. 13, pp. 189–222, Springer, Berlin, Germany, 2004.

[2] V. N. Sil'nikov and V. V. Vlassov, "Design of site-specific RNA-cleaving reagents," *Russian Chemical Reviews*, vol. 70, no. 6, pp. 491–508, 2001.

[3] H. Lönnberg, "Cleavage of RNA phosphodiester bonds by small molecular entities: a mechanistic insight," *Organic and Biomolecular Chemistry*, vol. 21, pp. 1687–1703, 2011.

[4] G. N. Grimm, A. S. Boutorine, and C. Hélène, "Rapid routes of synthesis of oligonucleotide conjugates from non-protected oligonucleotides and ligands possessing different nucleophilic or electrophilic functional groups," *Nucleosides, Nucleotides and Nucleic Acids*, vol. 19, no. 10–12, pp. 1943–1965, 2000.

[5] A. S. Modak, J. K. Gard, M. C. Merriman, K. A. Winkeler, J. K. Bashkin, and M. K. Stern, "Toward chemical ribonucleases 2. Synthesis and characterization of nucleoside-bipyridine conjugates. Hydrolytic cleavage of RNA by their copper(II) complexes," *Journal of the American Chemical Society*, vol. 113, no. 1, pp. 283–291, 1991.

[6] J. K. Bashkin, J. Xie, A. T. Daniher, U. Sampath, and J. L. F. Kao, "Building blocks for ribozyme mimics: conjugates of terpyridine and bipyridine with nucleosides," *Journal of Organic Chemistry*, vol. 61, no. 7, pp. 2314–2321, 1996.

[7] J. Hovinen, "Synthesis of carbon-3-substituted 1,5,9-triazacyclododecanes, RNA cleavage agents suitable for oligonucleotide tethering," *Bioconjugate Chemistry*, vol. 9, no. 1, pp. 132–136, 1998.

[8] H. Inoue, M. Shimizu, T. Furukawa, T. Tamura, M. Matsui, and E. Ohtsuka, "Site-specific RNA cleavage using terpyridine · Cu(II)-linked 2'-*O*- methyloligonucleotides," *Nucleosides and Nucleotides*, vol. 18, no. 6-7, pp. 1503–1505, 1999.

[9] J. K. Bashkin, E. I. Frolova, and U. Sampath, "Sequence-specific cleavage of HIV mRNA by a ribozyme mimic," *Journal of the American Chemical Society*, vol. 116, no. 13, pp. 5981–5982, 1994.

[10] T. H. Smith, J. V. LaTour, D. Bochkariov, G. Chaga, and P. S. Nelson, "Bifunctional phosphoramidite reagents for the introduction of histidyl and dihistidyl residues into oligonucleotides," *Bioconjugate Chemistry*, vol. 10, no. 4, pp. 647–652, 1999.

[11] D. M. Perrin, T. Garestier, and C. Hélène, "Bridging the gap between proteins and nucleic acids: a metal-independent RNAse A mimic with two protein-like functionalities," *Journal of the American Chemical Society*, vol. 123, no. 8, pp. 1556–1563, 2001.

[12] N. G. Beloglazova, N. N. Polushin, V. N. Sil'nikov, M. A. Zenkova, and V. V. Vlassov, "Site selective cleavage of yeast $tRNA^{Phe}$ by conjugates of antisense oligonucleotide with imidazole containing cleaving groups," *Doklady Academii Nauk*, vol. 369, no. 6, pp. 827–830, 1999.

[13] L. Yurchenko, V. Silnikov, T. Godovikova, G. Shishkin, J. J. Toulme, and V. Vlassov, "Cleavage of Leishmania mini-exon sequence by oligonucleotides conjugated to a diimidazole construction," *Nucleosides and Nucleotides*, vol. 16, no. 7–9, pp. 1721–1725, 1997.

[14] V. Vlassov, T. Abramova, T. Godovikova, R. Giege, and V. Silnikov, "Sequence-specific cleavage of yeast $tRNA^{Phe}$ with oligonucleotides conjugated to a diimidazole construct," *Antisense and Nucleic Acid Drug Development*, vol. 7, no. 1, pp. 39–42, 1997.

[15] K. Ushijima, H. Gouzu, K. Hosono et al., "Site-specific cleavage of tRNA by imidazole and/or primary amine groups bound at the 5′-end of oligodeoxyribonucleotides," *Biochimica et Biophysica Acta*, vol. 1379, no. 2, pp. 217–223, 1998.

[16] V. Silnikov, G. Zuber, J. P. Behr, R. Giege, and V. Vlassov, "Design of ribonuclease mimics for sequence specific cleavage of RNA," *Phosphorus, Sulfur and Silicon and Related Elements*, vol. 109-110, no. 1–4, pp. 277–280, 1996.

[17] J. Hovinen, A. Guzaev, E. Azhayeva, A. Azhayev, and H. Lönnberg, "Imidazole tethered oligodeoxyribonucleotides: synthesis and RNA cleaving activity," *Journal of Organic Chemistry*, vol. 60, no. 7, pp. 2205–2209, 1995.

[18] A. Y. Karyagin, T. V. Abramova, V. N. Sil'nikov, and G. V. Shishkin, "Reagents for addressed modification of biopolymers," *Russian Chemical Bulletin*, vol. 49, no. 3, pp. 540–545, 2000.

[19] M. A. Reynolds, T. A. Beck, P. B. Say et al., "Antisense oligonucleotides containing an internal, non-nucleotide-based linker promote site-specific cleavage of RNA," *Antisense and Nucleic Acid Drug Development*, vol. 24, no. 4, pp. 760–765, 1996.

[20] M. Endo, Y. Azuma, Y. Saga, A. Kuzuya, G. Kawai, and M. Komiyama, "Molecular design for a pinpoint RNA scission. Interposition of oligoamines between two DNA oligomers," *Journal of Organic Chemistry*, vol. 62, no. 4, pp. 846–852, 1997.

[21] N. N. Polushin, "The precursor strategy: terminus methoxyoxalamido modifiers for single and multiple functionalization of oligodeoxyribonucleotides," *Nucleic Acids Research*, vol. 28, no. 16, pp. 3125–3133, 2000.

[22] A. M. Morocho and N. N. Polushin, "Methoxyoxalamido chemistry in the synthesis of tethered phosphoramidites and functionalized oligonucleotides," *Current Protocols in Nucleic Acid Chemistry*, vol. chapter 4, p. unit 4.29, 2006.

[23] N. G. Beloglazova, N. N. Polushin, V. N. Sil'nikov, M. A. Zenkova, and V. V. Vlasov, "Site-specific cleavage of yeast tRNA(Phe) by derivatives of oligonucleotides bearing bisimidazole groups," *Doklady Akademii Nauk*, vol. 369, no. 6, pp. 827–830, 1999.

[24] N. G. Beloglazova, M. M. Fabani, M. A. Zenkova et al., "Sequence-specific artificial ribonucleases. I. Bis-imidazole-containing oligonucleotide conjugates prepared using precursor-based strategy," *Nucleic Acids Research*, vol. 32, no. 13, pp. 3887–3897, 2004.

[25] S. Nishikawa, H. Morioka, H. J. Kim et al., "Two histidine residues are essential for ribonuclease T1 activity as is the case for ribonuclease A," *Biochemistry*, vol. 26, no. 26, pp. 8620–8624, 1987.

[26] A. Wlodawer, M. Miller, and L. Sjölin, "Active site of RNase: neutron diffraction study of a complex with uridine vanadate, a transition-state analog," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 80, no. 12, pp. 3628–3631, 1983.

[27] L. Vitagliano, A. Merlino, A. Zagari, and L. Mazzarella, "Productive and nonproductive binding to ribonuclease A: X-ray structure of two complexes with uridylyl(2',5')guanosine," *Protein Science*, vol. 9, no. 6, pp. 1217–1225, 2000.

[28] M. Komiyama and T. Inokawa, "Selective hydrolysis of tRNA by ethylenediamine bound to a DNA oligomer," *Journal of Biochemistry*, vol. 116, no. 4, pp. 719–720, 1994.

[29] R. Giegé, B. Felden, M. A. Zenkova, V. N. Sil'nikov, and V. V. Vlassov, "Cleavage of RNA with synthetic ribonuclease mimics," *Methods in Enzymology*, vol. 318, pp. 147–165, 2000.

[30] M. Zenkova, N. Beloglazova, V. Sil'nikov, V. Vlassov, and R. Giegë, "RNA cleavage by 1,4-diazabicyclo[2.2.2]octane-imidazole conjugates," *Methods in Enzymology*, vol. 341, pp. 468–490, 2001.

[31] N. G. Beloglazova, V. N. Sil'Nikov, M. A. Zenkova, and V. V. Vlassov, "Cleavage of yeast $tRNA^{Phe}$ with complementary oligonucleotide conjugated to a small ribonuclease mimic," *Federation of European Biochemical Societies Letters*, vol. 481, no. 3, pp. 277–280, 2000.

[32] N. V. Tamkovich, A. N. Zenkov, V. V. Vlasov, and M. A. Zenkova, "An RNA sequence determines the speed of its cleavage by artificial ribonucleases," *Russian Journal of Bioorganic Chemistry*, vol. 36, no. 2, pp. 207–218, 2010.

[33] V. A. Petyuk, M. A. Zenkova, R. Giege, and V. V. Vlassov, "Hybridization of antisense oligonucleotides with the 3' part of $tRNA^{Phe}$," *Federation of European Biochemical Societies Letters*, vol. 444, no. 2-3, pp. 217–221, 1999.

[34] V. Petyuk, R. Serikov, V. Tolstikov et al., "Invasion of strongly binding oligonucleotides into tRNA structure," *Nucleosides, Nucleotides and Nucleic Acids*, vol. 19, no. 7, pp. 1145–1158, 2000.

[35] R. N. Serikov, V. A. Petyuk, V. V. Vlassov, and M. A. Zenkova, "Hybridization of antisense oligonucleotides with yeast $tRNA^{Phe}$: factors determining the efficiency of interaction," *Russian Chemical Bulletin*, vol. 51, no. 7, pp. 1156–1165, 2002.

[36] A. V. Vlasov, V. V. Vlasov, and R. Giege, "RNA hydrolysis catalyzed by imidazole as a reaction for studying the secondary structure of RNA and complexes of RNA with oligonucleotides," *Doklady Akademii Nauk*, vol. 349, no. 3, pp. 411–413, 1996.

[37] N. G. Beloglazova, A. Y. Epanchintsev, V. N. Sil'nikov, M. A. Zenkova, and V. V. Vlasov, "Highly effective site-selective cleavage of RNA by imidazoles-containing antisense oligonucleotide conjugates," *Molecular Biology*, vol. 36, no. 4, pp. 731–739, 2002.

[38] Y. Kumazawa, T. Yokogawa, H. Tsurui, K. Miura, and K. Watanabe, "Effect of the higher-order structure of tRNAs on the stability of hybrids with oligodeoxyribonucleotides: separation of tRNA by an efficient solution hybridization," *Nucleic Acids Research*, vol. 20, no. 9, pp. 2223–2232, 1992.

[39] I. Zegers, D. Maes, M. H. Dao-Thi, F. Poortmans, R. Palmer, and L. Wyns, "The structures of RNase A complexed with 3'-CMP and d(CpA): active site conformation and conserved water molecules," *Protein Science*, vol. 3, no. 12, pp. 2322–2339, 1994.

[40] C. F. Aguilar, P. J. Thomas, D. S. Moss, A. Mills, and R. A. Palmer, "Novel non-productively bound ribonuclease inhibitor complexes—high resolution X-ray refinement studies on the binding of RNAse-A to cytidylyl-2',5'-guanosine (2',5'CpG) and deoxycytidylyl-3',5'-guanosine (3',5' dCpdG)," *Biochimica et Biophysica Acta*, vol. 1118, no. 1, pp. 6–20, 1991.

[41] C. F. Aguilar, P. J. Thomas, A. Mills, D. S. Moss, and R. A. Palmer, "Newly observed binding mode in pancreatic ribonuclease," *Journal of Molecular Biology*, vol. 224, no. 1, pp. 265–267, 1992.

[42] J. C. Fontecilla-Camps, R. de Llorens, M. H. le Du, and C. M. Cuchillo, "Crystal structure of ribonuclease A·d(Ap TpApApG) complex. Direct evidence for extended substrate recognition," *Journal of Biological Chemistry*, vol. 269, no. 34, pp. 21526–21531, 1994.

[43] D. L. Birdsall and A. McPherson, "Crystal structure of ribonuclease A·d(ApTpApApG) complex. Direct evidence for extended substrate recognition," *Journal of Biological Chemistry*, vol. 267, no. 31, pp. 22230–22236, 1992.

[44] R. Kierzek, "Structural consideration concerning cleavage of RNA," in *Nucleic Acids and Molecular Biology: Artificial Nucleases*, M. A. Zenkova, Ed., vol. 13, pp. 33–48, Springer, Berlin, Germany, 2004.

[45] I. Zagórowska, S. Kuusela, and H. Lönnberg, "Metal ion-dependent hydrolysis of RNA phosphodiester bonds within hairpin loops. A comparative kinetic study on chimeric ribo/2'-*O*-methylribo oligonucleotides," *Nucleic Acids Research*, vol. 26, no. 14, pp. 3392–3396, 1998.

[46] I. Zagorowska, S. Mikkola, and H. Lönnberg, "Hydrolysis of phosphodiester bonds within RNA hairpin loops in buffer solutions: the effect of secondary structure on the inherent reactivity of RNA phosphodiester bonds," *Helvetica Chimica Acta*, vol. 82, no. 12, pp. 2105–2111, 1999.

[47] S. Mihkola, U. Kaukinen, and H. Lönnberg, "The effect of secondary structure on cleavage of the phosphodiester bonds of RNA," *Cell Biochemistry and Biophysics*, vol. 34, no. 1, pp. 95–119, 2001.

[48] U. Kaukinen, S. Lyytikäinen, S. Mikkola, and H. Lönnberg, "The reactivity of phosphodiester bonds within linear single-stranded oligoribonucleotides in strongly dependent on the base sequence," *Nucleic Acids Research*, vol. 30, no. 2, pp. 468–474, 2002.

[49] R. Häner, J. Hall, A. Pfützer, and D. Hüsken, "Development of artificial ribinucleases," *Pure and Applied Chemistry*, vol. 70, no. 1, pp. 111–116, 1998.

[50] R. Serikov, V. Petyuk, Y. Vorobijev et al., "Mechanism of antisense oligonucleotide interaction with natural RNAs," *Journal of Biomolecular Structure and Dynamics*, vol. 29, no. 1, pp. 27–50, 2011.

SAGE-Hindawi Access to Research
Journal of Nucleic Acids
Volume 2011, Article ID 870903, 15 pages
doi:10.4061/2011/870903

Review Article

A Potential of microRNAs for High-Content Screening

Andrius Serva, Christoph Claas, and Vytaute Starkuviene

BioQuant, University of Heidelberg, 69120 Heidelberg, Germany

Correspondence should be addressed to Vytaute Starkuviene, vytaute.starkuviene@bioquant.uni-heidelberg.de

Received 17 October 2010; Revised 15 May 2011; Accepted 3 June 2011

Academic Editor: Dmitry A. Stetsenko

In the last years miRNAs have increasingly been recognised as potent posttranscriptional regulators of gene expression. Possibly, miRNAs exert their action on virtually any biological process by simultaneous regulation of numerous genes. The importance of miRNA-based regulation in health and disease has inspired research to investigate diverse aspects of miRNA origin, biogenesis, and function. Despite the recent rapid accumulation of experimental data, and the emergence of functional models, the complexity of miRNA-based regulation is still far from being well understood. In particular, we lack comprehensive knowledge as to which cellular processes are regulated by which miRNAs, and, furthermore, how temporal and spatial interactions of miRNAs to their targets occur. Results from large-scale functional analyses have immense potential to address these questions. In this review, we discuss the latest progress in application of high-content and high-throughput functional analysis for the systematic elucidation of the biological roles of miRNAs.

1. Introduction

miRNAs (microRNAs) are 17-nt to 24-nt long noncoding RNAs that regulate gene expression in metazoans. miRNAs act by partially or completely complementary binding to their target mRNAs, resulting in translational repression and/or mRNA degradation [1, 2]. miRNAs are predicted to affect the expression of nearly 60% of protein-coding mammalian genes [3, 4] and, thereby, to control many, if not all, biological processes. Fundamental changes at the cellular and organismal level, including development [5], aging [6], the stress response [7], cell proliferation [8, 9], and apoptosis [10, 11], were shown to be regulated by miRNAs. Furthermore, miRNAs have been implicated in various diseases, such as diabetes [12–14], cancer [15, 16], hepatitis C [17], neurodevelopmental (reviewed in [18]), and mental [19] disorders. Rapidly growing knowledge of miRNAs as potent regulators in health and disease makes miRNAs attractive as targets for therapeutic intervention [20, 21] as well as for diagnostic markers [22, 23].

Numerous previous publications have addressed miRNA biogenesis and action (for detailed reviews see [24, 25]). Briefly, miRNAs are transcribed as long primary transcripts (pri-miRNAs), most of which are polyadenylated and capped. Pri-miRNAs are initially cleaved in nucleus by a multiprotein complex, called Microprocessor, yielding ~70-nt long stem-loop structured precursor miRNAs (pre-miRNAs). The key components of the Microprocessor complex are the RNase III enzyme Drosha and the double-stranded RNA-binding protein DGCR8/Pasha [26]. The excised pre-miRNA hairpin is then exported to the cytoplasm by Exportin-5 complexed with Ran-GTPase [27]. In the cytoplasm, the pre-miRNA is further processed to a 20–22-nt long miRNA/miRNA* duplex by a second RNase III enzyme, Dicer, which is in a complex with the TRBP and PACT proteins [28, 29]. Subsequently, the miRNA duplex is unwound by multiple helicases, which may be miRNA-specific and may regulate miRNA activity [30, 31]. The mRNA-targeting miRNA strand (guide strand) is loaded into the miRNA-induced silencing complex (miRISC). Until recently, it was assumed that the complementary miRNA* strand (passenger strand) is degraded, but there is now evidence that a substantial cohort of miRNA* species is functionally active [32]. The core components of miRISC are proteins of the Argonaute (AGO) [33] and GW182 protein families [34]. Individual miRNAs might need specific maturation steps [35–37].

Once incorporated into miRISC, the miRNA brings the complex to its target mRNAs by interacting with complementary binding sites, which can be present in multiple

copies [38–40]. Each miRNA can usually affect more than one transcript and, as a consequence, many proteins simultaneously [41, 42]. On the other hand, multiple miRNAs can repress expression of a single target mRNA [43–46]. miRNAs are postulated to preferentially bind to the 3′ untranslated regions (3′UTRs) of transcripts [47]. However, recent experimental evidences prove the existence of a new class of miRNA targets containing miRNA binding sites in both their 5′UTR and 3′UTR [48] or within the coding region of mRNA [49].

The complexity of miRNA-mediated modulation of gene expression is only beginning to be appreciated, and much research needs to be done in order to understand miRNA global and adaptive regulatory functions. In this review, we summarize available methodologies for modulating expression levels of endogenous miRNAs, as well as on the application of these strategies for high-content and high-throughput functional studies.

2. RNA Silencing

The discovery that small ncRNAs (noncoding RNAs) play pivotal roles in fundamental biological processes has considerably widened our knowledge of mechanisms of gene regulation in the last years [50, 51]. siRNAs (small interfering RNAs) and miRNAs are the two best characterized classes of ncRNAs. Both are derived from dsRNA (double-stranded RNA) precursors and exert their inhibitory function on gene expression by Watson-Crick base-pairing to complementary sequences in target RNA molecules: an effect commonly referred to as gene silencing [2]. Moreover, both siRNAs and miRNAs share some components of the cellular effector machinery involved in gene silencing [52]. Usually, miRNAs modify expression of endogenous genes whereas siRNAs have evolved to defend genome integrity against foreign invaders, like viruses or transposons [53]. In mammals siRNAs are 21-22 nt long fragments often, but not always, derived from foreign dsRNA by Dicer [52, 54]. siRNAs are incorporated into the siRNA-induced silencing complex (siRISC) [55] and bind to perfectly matched sequences in target molecules. Typically, this induces degradation of the bound RNA, a function called RNAi (RNA interference) [53], and this property was widely implicated in functional studies over the last decade (see what follows).

Binding of miRNA to perfectly matched sequences in mRNA can also result in degradation of the mRNA [56], but usually, miRNAs bind to sites in mRNAs with only partial sequence complementarity. This results primarily in translational repression rather than degradation [57] but can also cause secondary nucleolytic degradation of the mRNA [24]. Conversely, binding of siRNAs to partly unmatched sequences in mRNAs can result in translational blockade-potentially interfering with the results from siRNA-based screening experiments [56].

3. siRNA/shRNA-Based Screens

In the last years RNAi has developed into a powerful tool for systematic studies of fundamental physiological and pathological processes. A number of large-scale screens have been completed, analysing diverse cellular processes. Recent impressive examples of genome-wide RNAi-based screens in human cells are provided by the work of Collinet et al. [58], who performed a high-content survey of genes involved in endocytosis, and by the study of Neumann et al., who used time-lapse microscopy in living cells to identify genes that play a role in cell division [59].

siRNA-based screens have also been performed in several model organisms. *C. elegans* and Drosophila are especially receptive to this type of genetic screening [60]. In *C. elegans*, for example, gene silencing can be accomplished by feeding bacteria that express long dsRNA (about 200–2000-bp-long) or by providing such dsRNA in the medium [61, 62]. RNAi in cultured cells of Drosophila can similarly be induced by adding *in vitro* transcribed dsRNA to the culture medium [63]. In mammalian cells, in contrast, short RNA duplexes of 21–29 bp have to be administered, because long dsRNA evokes a response of the innate immune system, which ultimately leads to apoptosis [64]. The short siRNAs can be chemically synthesized or derived from transcribed PCR products by digestion with recombinant Dicer or bacterial RNase III (esiRNAs-endoribonuclease-prepared short interfering RNAs) [64, 65]. siRNA-mediated knock-down can attain close to 100% reduction in the target mRNA. However, the effectiveness of an individual siRNA is hard to predict and, therefore, several siRNAs targeting different regions of the target mRNA have to be tested. Some authors suggest checking 4–6 siRNAs per gene to obtain reliable results in a screening experiment [66].

If stable knock-down is needed or if cells are difficult to transfect, short hairpin RNAs (shRNAs) can be used. The sequences encoding the shRNA can be cloned into plasmids or virus-derived vectors (lenti-, retro-, or adenoviral origin) [67–69]. In the case of "second generation" shRNAs, so called shRNA-miRs, the RNAi-triggering small RNA sequence is cloned into the backbone of a pri-miRNA [70]. This design principle, together with improved selection of the small RNA targeting sequence, improves both the production levels of small RNAs and the silencing efficiency. shRNA-miRs can be cloned into constructs carrying different promoters and used for tissue-specific or inducible expression [70].

4. Modulation of miRNA Function as an Approach for Functional Screening

Clearly, their biological importance *per se* makes analyses of miRNA expression and identification of their target mRNAs a focus of the current research. Yet specific features of miRNAs give them strong potential as tools in functional genomics. Firstly, miRNAs are able to change the translation of hundreds of mRNAs simultaneously and, by doing so, add another layer of regulation to gene expression [71]. In fact, miRNAs can influence the whole biological programs, including development, apoptosis, proliferation, and differentiation, not only through direct interactions with target mRNAs but also indirectly by altering expression of, for example, components of the translation or RNAi machinery [42].

In general, miRNAs effect only subtle modulation of target gene expression. Their pronounced effect on cellular behaviour might, therefore, be a consequence of their ability to influence multiple genes involved in a single pathway. Indeed, miRNA-based screens can provide lists enriched in functionally related targets: for example, a recent RNAi screen identified three target genes that phenocopy miR-19 and cooperate in the regulation of phosphatidylinositol-3 kinase-mediated survival signalling [72]. miRNA-based screens and, especially, a combination of miRNA- and siRNA-based screens, though laborious, might have a clear advantage over siRNA-based screens alone, since with siRNAs, downregulation of only single individual mRNAs is expected. On the other hand, enrichment of target genes with specific functions is not always found using current bioinformatic prediction tools [73]; therefore, the potential to organise hits of miRNA-based screens directly into functional networks is as yet unproven. The fact that a single mRNA can be inhibited by different miRNAs [74] also means that cellular phenotypes result from synergistic effects, complicating interpretation. Having an infrastructure for large-scale experimentation, complex and previously unanticipated regulatory patterns might be identified.

In contrast to the vast number of mRNAs that have to be knocked-down in comprehensive functional genomic approaches using siRNAs/shRNAs, the number of miRNAs that have to be analysed in a genome-wide screen is relatively low. To date, 1424 miRNAs have been identified in the human genome (miRBase database, release 17; http://www.mirbase.org/). Various *in silico* methods, developed to predict targets of miRNAs, estimate that between 10% and 60% of all mRNAs might be influenced by miRNAs [3, 4, 75, 76]. Hence, the analysis of the comparatively small number of miRNAs could, in principle, cover the function of substantial fraction of human genes. To modulate the expression of the same number of genes by siRNA/shRNA would require about 44,000 to 66,000 different siRNA/shRNAs, taking into account the need for multiple reagents targeting the same mRNA (see above).

5. Tools to Modulate miRNA Function

Here, we will briefly describe approaches that are currently in use to modulate miRNA function and discuss their potential and limitations (Table 1). Basically, it is possible either to interfere with miRNA expression (loss-of-function assays) or to induce ectopic (over-) expression of miRNA (gain-of-function).

One possibility for analysis of miRNA would be a complete knock-out. This very precise intervention will result in a complete loss of function. The resulting phenotypes are, therefore, often stronger than those seen after knock-down. For some time, generation of knock-out phenotypes was feasible in mammals only by exploiting *homologous recombination* technology in mice [77]. More recently, germline competent embryonic stem cells have also been established in rat, so that homologous recombination can now be performed in this species as well [78]. The development of genome editing approaches, using engineered *zinc finger nucleases*, opened up the possibility to extend such analyses to other organisms and tissue culture cells [79]. Still, a gene knock-out is laborious and, if the work is given to a contractor, expensive. Many miRNAs are located in introns of protein coding genes; so knock-out of a miRNA can result in simultaneous inactivation of the "host" protein coding gene with potentially detrimental effects that are unrelated to miRNA action. Finally, as discussed below, it is often desirable to downregulate the function of several miRNAs together, which is not feasible through knock-out. Interference with multiple miRNAs that are not synthesized from a single polycistronic gene is much easier to achieve using a knock-down approach.

The knock-down of miRNAs can be performed by application of *antisense oligonucleotides (ASOs)*. Usually, ASOs exert their effects independently of the cellular silencing machinery: their potential in biological assays was described as early as 1978 [80]. Since then, lot of experience has been gained in the design of stable, specific, and potent antisense reagents for research and therapeutic use. Since unmodified oligonucleotides are quickly degraded by nucleases when administered to cells, chemical modification is necessary to enhance stability and potency. (For a review on chemical modification used in antisense approaches, see [81].) Inactivation of complementary RNA can be achieved, for example, by delivery of DNA oligonucleotides. RNA/DNA hybrids are then recognized and cleaved by RNase H [82]. Alternatively, oligonucleotides can be coupled to ribozymes or DNAzymes that inactivate bound RNA by cleavage [81]. For the inactivation of miRNAs, RNA-based oligonucleotides that do not carry enzymatic activity have gained popularity. It is not clear whether these oligonucleotides work by inducing degradation of the targeted miRNA [83] or by forming stable ASO/miRNA heteroduplexes and, thereby, blocking miRNA function [84, 85].

ASOs and ribozymes that interfere with miRNA function are commonly referred to as antagomiRs and antagomiRzymes, respectively, and can be obtained from commercial sources. In addition, plasmid- and virus-derived vectors exist that both allow for simultaneous and permanent expression of antogomiRs targeting different miRNAs and permit measurement of expression levels by coexpression of a fluorescent reporter protein [86].

In principle, miRNA function can be inhibited at different stages [87]. For example, ASOs have been used to interfere with the maturation of pri-miRNA [88]. Targeted degradation of pri-miRNAs in the nucleus with RNase H-based ASOs has also been tried [87]. Strategies to target pri-miRNAs could have the advantage of combined inhibition of several miRNAs transcribed from the same polycistronic gene locus. Experiments with siRNAs targeting the loop region of pre-miRNA have been reported [89], but the approach has not acquired much popularity: the loop region might be difficult to access, leading to inefficient knock-down of miRNA [87].

With regard to knock-down of protein-coding mRNAs, siRNA are nowadays in much wider use than ASOs, probably because siRNAs provide a more potent and efficient mode

Table 1: Ways to analyse function of miRNA.

Aim	Method	Features	
		Pros	Cons
Knock-out of miRNA gene	Homologous recombination/Gene editing with zinc finger nucleases	(i) Precise intervention (ii) Complete loss-of-function	(i) Laborious and time consuming (ii) Simultaneous knock-out of protein encoded by the same transcriptional unit
Knock-down of miRNA	Antisense oligonucleotide Ribozymes/DNAzymes	(i) Can attack mature miRNA and miRNA precursors (ii) Stable expression from vector-constructs (iii) Commercially available (iv) Transfection/transduction of multiple oligonucleotides is possible	(i) OTEs and unspecific secondary effects may occur (ii) Efficiency of knock-down is difficult to predict (iii) Multiple knock-downs might be needed to produce a phenotype
	miRNA sponges	(i) Easy design (ii) Simultaneous knock-down of multiple miRNAs (iii) Expression can be verified by fluorescent reporter	(i) Efficiency of knock-down is difficult to predict (ii) Incomplete knock-down of individual miRNAs
Over-expression of miRNA	miRNA mimics	(i) Efficient silencing of target mRNA (ii) Oligonucleotides mimicking mature miRNAs or miRNA precursors could be used (iii) Transfection/transduction of multiple mimics is possible (iv) Stable transfection from vector-constructs	(i) Oversaturation of RNAi machinery can lead to secondary effects
	Conditional release of miRNA from riboswitch constructs	(i) Controlled expression of mature miRNA (ii) Possibly less toxic side effects	(i) Laborious design (ii) Lack of availability (iii) Not yet adaptable to high-throughput analysis
Release of target mRNA from repression by miRNA	Target protectors	(i) Release of specific mRNA from regulation by miRNA possible	(i) Target mRNA has to be known

of intervention. In some model organisms, like zebrafish, however, a certain class of ASOs, called "morpholinos", is still indispensable for knock-down of gene expression and so far could not be replaced by siRNA-based approaches [90].

More recently, vector constructs have been devised that contain multiple binding sites for miRNAs, in order to compete with endogenous mRNAs for miRNA binding. Such *miRNA sponges* appeared to be a versatile tool for miRNA research [91]. Loya and colleagues analysed phenotypes when miR-7, miR-8, and miR-9a in Drosophila were down-regulated by miRNA sponges. The authors found that miRNA sponges evoked the same phenotype as produced by homologous recombination, but in a milder form. This is expected: miRNA sponges do not completely sequester target miRNAs, because of the competitive binding of miRNAs to mRNAs. Still, this approach enables not only simultaneous down-regulation of several miRNAs but also inducible expression of miRNA sponges in specific tissues and at different time points, allowing spatiotemporal analysis of miRNA function. For instance, it was possible to assign the known role of miR-8 in neuromuscular junction formation by expression of miRNA sponge that targeted miR-8 in muscle cells [92].

As an alternative to, or to complement, loss-of-function studies, *miRNA-mimics* can be used to achieve *de novo* or enhanced expression of miRNA (gain-of-function). miRNA-mimics can be administered as synthetic small dsRNA with sequences identical to those of endogenous miRNAs [93, 94]. They resemble siRNA molecules and are readily incorporated into the cellular RNA silencing machinery. If stable ectopic expression of miRNA is desired, constructs are at hand that allow expression of single or multiple miRNAs, and measurements of miRNA expression level by cotranscription of fluorescent reporter protein are possible (e.g., [95]). As for other vector-based constructs, the choice of suitable promotors can render tissue- and/or time-specific expression of miRNA-mimics possible. Importantly, molecules mimicking different maturation stages of endogenous miRNAs can be overexpressed and analysed.

Spatial and temporal regulation of miRNA expression can also be attained by allosteric ribozymes, *riboswitches* [96]. These are modular constructs containing an aptamer domain (an RNA sequence that specifically binds a chemical compound) embedded within a ribozyme and fused to a pri-miRNA analogue. This construct is functionally inactive in the absence of the appropriate chemical trigger. Upon application of the trigger substance (e.g., theophylline in the case of the cited work) a conformational change is induced in the ribozyme leading to its activation. The ribozyme will then cleave the modular RNA in *cis,* with pri-miRNA released from the construct and processed [96]. The large number of known aptamers and their endogenous and exogenous chemical triggers could develop this approach of *conditional RNA interference* into a versatile alternative for inducible vector constructs, especially if toxic side effects in knock-down or knock-out approaches are of concern [97].

Finally, specific mRNAs can be released from miRNA-mediated inhibition by application of *target protectors*. In experiments performed by Choi et al. in zebrafish [98], morpholino-based ASOs were used, which prevented miR-430 binding by blocking seed-matched sites and neighbouring nucleotides in two target mRNAs. In this way, the authors succeeded in specifically interfering with miR-430-mediated translational repression. To conduct this type of experiment, it is necessary to know the targets of the miRNA of interest, but target protectors can be valuable in target validation and in characterization of miRNA: mRNA interactions.

6. miRNA-Based Functional Screens

The established infrastructure for siRNA/shRNA screens (i.e., robotics for large-scale sample preparation, automated data acquisition and analysis, data storage capacities) can easily be applied for high-throughput studies of miRNA function. As for siRNA-based screens, lipid-based transfection is most commonly used to achieve a transient overexpression [99] or inhibition [100] of miRNAs in cell culture. A number of screens were completed under conditions of miRNA stable overexpression [40, 45, 101, 102], achieved by transduction with retroviral vectors encoding specific miRNA genes [40]. Alternatively, constitutive overexpression or down-regulation of non-coding RNAs can be obtained by adenoviral vector-based systems [69, 103]. The advantages of the reverse transfection method, coupled with an automatic liquid handling system, have been also utilized in miRNA-based screens by several groups with a high success rate [104, 105].

The first miRNA was described in the early 1990s [50], but the initial accumulation of data concerning regulatory roles was rather slow [106–108]. This in turn delayed the availability of reagents to modulate the activity of endogenous miRNAs, and so the first functional miRNA-based screen [100] was completed a couple of years later than siRNA-based screens [109]. As a consequence, though, miRNA-based screens could benefit a lot from prior experience accumulated with siRNA-based screens.

7. Biological Processes Analysed in miRNA-Based Screens

A number of miRNA-based screens have been completed during the last 5 years (Table 2). Two major groups of biological processes have been investigated so far, namely, (i) cell viability, proliferation, and apoptosis and (ii) gene transcription and/or activity regulation.

Numerous miRNAs were shown to inhibit (let-7, miR-34a, miR-143, miR-145, miR-221, miR-222) [117–121] or to stimulate (miR-21, miR-133) [122, 123] cell proliferation. Moreover, recent studies demonstrated that a well-studied cluster of miR-17-92 can act both as oncogenes [124–127] and as tumour suppressors [9, 128], and these roles are probably cell type- and/or environment-dependent. Context-dependent activity was also reported for miR-24: down-regulation of miR-24 inhibited proliferation of A549 cells but increased growth of HeLa cells [100].

A considerable increase in the numbers of miRNAs that regulate cell proliferation and apoptosis was obtained

Table 2: miRNA-based functional screens.

Model system	Number of miRNAs screened	Type of regulation	Phenotype measured	Assay	Year	Reference
HeLa and A549	95	Loss-of-function by miRNA inhibitors	Cell proliferation and apoptosis	Cell counting and caspase-3/7 activity assay	2005	[100]
Primary hTERT-immortalized BJ-EHT fibroblasts	~450	Gain-of-function by stable miRNA expression	Sustained proliferation	miR-Array	2006	[40]
HeLa	~450	Gain-of-function by stable miRNA expression	$p27^{Kip1}$ regulation	GFP reporter assay and miR-Array	2007	[45]
MDA-MB-453	187	Gain-of-function by miRNA precursors	TRAIL-induced caspase-3 activation	Caspase-3/7 activity assay	2007	[104]
Drosophila clone 8 cells	77/78 miRNA loci	Gain-of-function by plasmid-based miRNA expression	Wg signaling pathway regulation	Luciferase reporter assay	2007	[110]
Neuroblastoma cell lines	8	Gain- and loss-of-function by miRNA mimics and inhibitors	Proliferation	Change in electrical impedance	2008	[111]
MCF7	~450	Gain-of-function by stable miRNA expression	Cell migration	Trans-well cell migration assay	2008	[101]
HeLa	91	Gain-of-function by miRNA mimics	p53 gene activity	Luciferase reporter assay	2009	[38]
HeLa	~450	Gain-of-function by stable miRNA expression	*Per* gene family regulation	GFP reporter assay and miR-Array	2009	[102]
Primary ovarian granulosa cells	80	Gain-of-function by miRNA precursors	Progesterone, testosterone and estradiol release	Enzyme immunoassay (EIA)	2009	[112]
HEK 293	266	Gain-of-function by miRNA mimics	p21Cip/Waf1 regulation	Luciferase reporter assay	2010	[46]
HCT-16	810	Gain-of-function by miRNA mimics	Cell viability in the presence of Bcl-2 family inhibitor ABT-263	CellTiter-Glo Luminescent Cell Viability Assay	2010	[99]
HEK 293T	107	Gain-of-function by miRNA mimics	p53 gene regulation	Luciferase reporter assay	2010	[39]
HCT116 $p53^{+/+}$, H460 and MCF7	5	Gain-of-function by miRNA precursors	p53 gene regulation	Western blot	2010	[10]
Huh-7	327	Gain-of-function by miRNA precursors	Lipid droplet formation and growth	Immunocytochemistry and fluorescence microscopy	2010	[105]
Primary ovarian granulosa cells	80	Gain-of-function by miRNA precursors	Proliferation and apoptosis	Immunocytochemistry and fluorescence microscopy	2010	[113]
DLD-1	319	Gain-of-function by miRNA precursors	Cell viability	CellTiter-Glo Luminescent Cell Viability Assay	2010	[11]
HMEC	328	Gain-of-function by miRNA precursors	Proliferation	Fluorescence microscopy	2010	[114]

Table 2: Continued.

Model system	Number of miRNAs screened	Type of regulation	Phenotype measured	Assay	Year	Reference
MIA PaCa-2	445	Gain-of-function by stable miRNA expression	Proliferation	Custom-made microarray	2010	[115]
HeLa and HeLa P4	8	Gain- and loss-of-function by miRNA precursors and inhibitors	Cell proliferation and trafficking	Fluorescence microscopy	2011	[116]

[40, 99, 105] since Cheng et al. reported the first large-scale screen to identify mammalian miRNAs involved in these processes [100]. In 2007, Ovcharenko et al. performed a screen with 187 synthetic oligonucleotides to capture the modulators of TRAIL-induced apoptotic pathway [104]. They found that 34 of tested miRNAs modulate the activity of caspase-3. Recently, a gain-of-function miRNA-based screen was completed in human colorectal cancer DLD-1 cells [11]. By measuring cell viability, authors not only confirmed function of already known oncogenic miR-372 and miR-373 [40] but also discovered novel miRNAs involved in cell proliferation and apoptosis. miR-491 was among the strongest antiproliferative miRNAs, and further experimental analysis revealed that it induces apoptosis via direct down-regulation of antiapoptotic Bcl-xL [129]. miRNAs regulating expression of another member of Bcl-2 protein family, Mcl1, were identified by screening a library of 810 human miRNAs for the ability to confer resistance of cancer cells to ABT-263, an inhibitor of Bcl-2 family members [99]. Viability measurements of human colorectal cancer HCT-16 cells transiently transfected with miRNA mimics revealed 19 miRNAs that sensitized cells to ABT-263. 15 of these miRNAs showed the same phenotype in melanoma CHL1 cells. Furthermore, 10 out of 12 strong sensitizer miRNAs that were analysed for targeting 3′URT of Mcl1 were confirmed as direct regulators of the gene. These examples of the screens demonstrate that miRNAs that modulate sensitivity to chemotherapeutic agents can be identified and potentially, in future, used in cancer therapy [21, 130]. The feasibility of such screens has vastly improved by the evolution of methods to quantify cell proliferation from straightforward cell counting [100] to recording of electrical impedance over 96 hours [111].

miRNAs acting on transcription and/or gene activity regulation are usually identified in so-called "target-based screening". In the most cases, luciferase or green fluorescent protein (GFP) is fused to the gene of interest and the strength of the detected signal is used to gauge the expression or activity of the test gene [38, 39, 45, 46, 102]. Using a functional genetic approach with stable expression of individual miRNAs [40], miR-221 and miR-222 were demonstrated to specifically regulate expression of tumour suppressor $p27^{Kip1}$ [45], miR-192/194 cluster-modulate expression of Per gene family [102].

The repertoire of cellular processes analysed by miRNA functional screens is expanding rapidly (Table 2). Doing miRNA-based screens in appropriate cellular contexts, for instance, for miRNAs regulating steroidogenesis in ovarian cells [112], and for miRNAs regulating lipid droplet formation in hepatocytes [105] helps to ensure acquisition of physiologically relevant information.

8. Gain-of-Function versus Loss-of-Function in miRNA-Based Screens

Although reagents for gain-of-function and loss-of-function miRNA experiments are available to similar extents, nearly all reported screens to assess the function of miRNAs utilized the gain-of-function approach (Table 2). This trend probably arose because of the ease with which ectopic expression of miRNAs can be confirmed. For instance, exogenous reporter gene assays [116, 131], qRT-PCR, and Northern blotting and ribonuclease protection assay [40] are easily applicable methods to measure over-expression of miRNAs. In contrast, the only assays that have been extensively used so far to show the inhibition of endogenous miRNA activity have involved reporters [38, 100, 128, 132]. The popularity of gain-of-function screens might additionally be explained by potentially easier evaluation. Over-expression of miRNAs might induce accentuated phenotypes, which might be not related to the levels of endogenous miRNAs, whereas the evaluation of data obtained under the conditions of miRNA down-regulation is possible only then one knows expression level of endogenous miRNAs in the test system. Acquisition of these data sets, therefore, needs thorough additional experimentation.

9. Fluorescence Screening Microscopy for miRNA-Based Screens

There are virtually no reasons why the read-out strategies in miRNA-based screens should be different from the ones established in siRNA-based screens. Nevertheless, we consider fluorescence microscopy screens to be highly advantageous. Features that make such microscopy ideal to analyse regulatory potential of miRNAs include the following: (i) rapid collection of large amount of data, (ii) feasibility of phenotype multiplexing, (iii) the possibility to acquire quantitative data on a cell-by-cell basis and/or population-based basis, and (iv) detection of subtle phenotypes [133, 134]. One of the first studies applying this

technology for functional miRNA studies was by Sirotkin and colleagues, who reported an immunocytochemistry- and fluorescence-microscopy-based screen to identify miRNAs regulating proliferation and apoptosis [113]. Primary human ovarian cells were transfected with synthetic miRNA precursors, and PCNA and cyclin B1 proteins were used as markers of proliferation. The expression levels of Bax and caspase-3, in combination with the TUNEL assay, were used to determine the extent of apoptosis. The power of technology is demonstrated by another microscopy-based screen: using automated image analysis and nucleus classification software, a novel antiproliferative activity of miR-320a was discovered [116]. The screen to identify miRNAs controlling lipid droplet formation in hepatocytes [105] illustrated the sensitivity of fluorescence microscopy-based approach, which was comparable to the laborious biochemical assay—11 out of 327 transiently overexpressed human miRNAs were selected via an automated work flow as the most potent regulators of intracellular lipid content [105].

10. Validation of Hits in miRNA-Based Screens

In current screens, the list of primary hits involved in regulation of a biological pathway can be obtained reasonably fast, if necessary infrastructure is on place. The primary hit lists in siRNA/shRNA-based screens are usually validated by (i) repeating the assay with different types of reagents, and (ii) secondary assays, and (iii) rescue experiments, in which a phenotype caused by knock-down of a certain gene product is rescued by the over-expression of a construct that cannot be attacked by siRNA (either due to mutation or because of origin from another species). Currently, there is no clear agreement on criteria for successful validation of primary hit miRNAs, except from reproducing the phenotype with different reagents. Additionally, the effects caused by miRNA over-expression and down-regulation could be compared. Since this can be laborious, lists of miRNAs that cause a particular phenotype in screens may usefully be published without further analysis [100, 112, 113], allowing follow-up by others.

In order to perform a follow-up research on hit miRNAs, prediction of their target genes needs to be done first. Many computational algorithms have been developed over the last decade to predict miRNA targets [135], and a combination of multiple algorithms is frequently used to narrow down the candidates. Interesting strategy was used in the studies of TRAIL-induced apoptosis when hits of miRNA-based screen were compared with those of siRNA-based screen in order to identify plausible interactions between miRNAs and their target mRNAs [104].

Many of the screens that have been completed so far have been extended to experimental identification and validation of the targets of the hit miRNAs (Table 3). The most widely used method to test a direct regulation of gene expression by miRNAs is based on reporter assays. The 3′UTR, the 5′UTR, or the whole gene is cloned immediately downstream of a reporter gene encoding luciferase or fluorescent protein. The construct is then transiently cotransfected with miRNA mimics or antimiRs into host cells and luciferase activity or fluorescence is measured after 24–48 hours of incubation [10, 11, 38, 99, 110]. Additional evidence that the target gene is directly regulated by an miRNA can be provided by mutating [38, 110] or deleting [10, 101] predicted miRNA binding sites in the 3′UTR of a reporter vector. As controls, the entire 3′UTR can be inserted in the reverse orientation [136] or truncated [110]. Frequently, qRT-PCR has been used to show degradation of target mRNAs in miRNA overexpressing cells [10, 45, 46, 101].

Since miRNAs can execute their regulatory action by repressing translation and/or promoting mRNA decay of target genes [137], measurement of the product protein level is an additional means to validate miRNA action [11, 45, 46]. Biochemical techniques for the isolation of target mRNAs were also shown to be rewarding [138–140]. Eventually, the action of miRNA can be tested by down-regulation of known or putative target genes and monitoring the resulting phenotypes [72, 105]. Observation of the same phenotype after both miRNA over-expression and down-regulation of its putative target genes strengthens evidence that the target identification was valid. Such tests might include down-regulation of a single target as well as multiple ones. Complex evaluations are possible if combining several validation approaches. For example, changes in target protein level without any changes in a reporter assay would suggest that the miRNA does not directly target an affected gene but rather regulates its modulator or the whole pathway [99].

11. Challenges of siRNA-, shRNAs-, and miRNA-Based Screens

RNAi-based screens have turned to be extremely useful for studies in various fields of biology and medicine [59, 141–143]. A robust assay, careful design of the experiment, reliable controls, and exhaustive testing of the reagents are crucial for successfully completing a screen, and several excellent reviews, describing the planning of RNAi screen experiments, have been published (e.g., [60, 144–146]). We will now focus on some practical aspects of siRNA/shRNA-based screens that could be relevant while performing miRNA-based screens.

The analysis of siRNA-based screens is often complicated by off-target effects (OTEs): phenotypes caused by unspecific down-regulation of mRNAs. One source of OTEs is partial homology of siRNAs with unintended mRNAs. A different type of OTE arises from induction of cytokine production by siRNA/shRNA with certain sequence motifs [147], which ultimately results in apoptosis. Therefore, in the last years considerable effort has been invested in improving the performance of siRNAs/shRNAs. For example, induction of an interferon response can be mitigated by excluding GU-rich sequence motifs in siRNA molecules [148]. Specificity can be considerably enhanced by altering the chemical backbone of siRNAs: examples are LNA (locked nucleic acid) nucleotide analogues or modifying ribosyl groups by addition of a 2′-O-methyl group at specific positions [149, 150].

Table 3: Validation assays in miRNA-based screens.

Effector miRNA	Target genes	Validation assay					Year	Reference
		Expression[a]	Reporter[b]	qRT-PCR	IS[c]	siRNA[d]		
miR-372 and -373	LATS2	+	+	+	+		2006	[40]
miR-221 and -222	$p27^{Kip1}$		+		+		2007	[45]
miR-315	Axin and Notum		+				2007	[110]
miR-34[a]	Bcl2 and MYCN				+		2008	[111]
miR-373 and -520c	CD44	+	+				2008	[101]
miR-29[a,b,c]	p85α and CDC42		+		+		2009	[38]
miR-192 and -194	Per1, 2 and 3		+				2009	[102]
28 miRNAs	p21Cip/Waf1		+		+		2010	[46]
10 miRNAs	MCL1		+				2010	[99]
miR-1285	p53		+	+	+		2010	[39]
miR-504	p53		+		+		2010	[10]
11 miRNAs	Multiple genes					+	2010	[105]
miR-491	Bcl-xL		+	+	+		2010	[11]
28 miRNAs	p21Cip/Waf1		+		+		2010	[114]

[a] mRNA expression profiling.
[b] Luciferase and/or GFP reporter assays.
[c] Immunostaining.
[d] Recapitulation of miRNA-mediated phenotype by selected siRNAs.

Secondary or even toxic effects caused by transfection reagents alone and/or in combination with silencing RNA cannot be completely ruled out. Fortunately, current formulations contain 5 to 10 times less reagents than were required several years ago. Another approach is to pool several siRNAs against the same target, while using reduced amounts of each. The idea behind this approach is that the target mRNA will be attacked by all, or the majority of, siRNAs in the pool, and the specific down-regulations will act cumulatively, whereas OTEs of individual siRNAs will be too weak to affect results. The same concept of improved signal-to-noise ratio in observed phenotypes forms also the basis for the application of esiRNAs [65].

Despite these strategies, though, OTEs and unspecific secondary effects cannot be avoided completely. It is, therefore, essential to control for both. One option is to use several different siRNAs to target the same mRNA at independent sites. The probability that an observed phenotype corresponds to an on-target silencing effect rises with the number of siRNAs causing the same phenotype [151]. Moreover, it is of advantage to measure several cellular parameters (e.g., cell shape and size, number of cells) in a screen. Multiparametric analysis allows generation phenotypic profiles of individual siRNAs. A high degree of overlap in these profiles justifies strong confidence in the specificity of an observed phenotype [58, 151].

One of the most serious concerns in functional screens is possible oversaturation of the cellular silencing machinery by exogenous siRNA/shRNA and miRNA mimics. This can result in derepression of mRNAs that are regulated by endogenous miRNAs [152, 153]. Various strategies have been developed to deal with these problems (for a detailed compilation of strategies see [154]). Briefly, overloading the cellular RNA silencing machinery can be avoided by applying the lowest possible amounts of siRNA/shRNA to the cells. For instance, transcription from RNA Pol II-driven promotors, and use of inducible constructs, can be useful to control levels of vector-encoded shRNAs [70, 155]. Using vectors derived from adeno-associated virus can also be a way to induce a moderate level of stable shRNA expression in cells [21, 156]. Coexpression of AGO-2 has also been described as an effective approach to overcome the problem of saturating the RNAi machinery by ectopic siRNA, shRNA, and miRNA [157]. Enhanced amounts of AGO-2 were shown to drive RNAi interference toward preferential knock-down of perfectly matched target mRNAs in diverse mammalian cell lines. Because, in addition, less siRNA/shRNA was needed for specific knock-down, and unspecific targeting of mRNAs was significantly reduced in the presence of ectopic AGO-2, the authors suggest that this strategy might result in reduced rates of false negatives and false positives in RNAi-based screening approaches. On the other hand, the suitability of the approach should be tested rigorously, as AGO-2 was recently shown to play a role in biogenesis of particular type of miRNAs [36, 37].

Finally, miRNA-based functional screens might suffer from neutral repression, when reduced transcript levels do not lead to measurable changes in cellular behaviour. Sometimes, residual protein is still sufficient to fulfil the required function; alternatively, feedback mechanisms exist that compensate functionally for the knock-down [158].

Usually, screens encompass thousands of genes or even the whole genomes; so it is hardly possible to look for functions that might be conferred by redundant gene function. Often, cellular functions are not controlled by just one protein, but by two or more isotypes or evolutionary-related forms of proteins to ensure higher reliability of the process, and metabolic pathways may be redundant.

The need to use several siRNAs per protein, and multiple possible combinations of distinct siRNAs, limits studies that aim to identify redundant gene products to relatively small, directed screens. Conversely, highly homologous gene products, like, for example, splice variants, might have distinct cellular functions, but siRNAs might target all forms together (e.g., [159]). Similar concerns pertain to miRNA-based screens, as redundant activities of miRNAs in regulation of many biological processes is well documented [160]: Miska et al. reported that most miRNAs in *C. elegans* are individually dispensable [161], and Voorhoeve et al. found that miR-372 and miR-373 cooperate with oncogenic RAS in the development of testicular germ cell tumours [40]. A way to deal with miRNA redundancy could be the use of sensitized genetic backgrounds as described by Brenner et al. [162]. By deleting one of the two AGO genes in *C. elegans* the authors partially disabled the RNA silencing machinery. In addition, worms with defects in chromatin modification or transcriptional regulation were generated. Using these organisms allowed definition of biological roles for several individual miRNAs [162]. Additionally, synthetic phenotypes might be used to analyse functions of miRNAs [160].

The complexity of miRNA action presents a challenging task for high-throughput functional analysis. However, given our experience from RNAi screens over the last decade, with siRNAs and shRNAs in different model organisms, careful experimental design and exhaustive target validation makes this powerful technology indispensable for understanding the biological roles of miRNAs.

Acknowledgments

The authors are grateful to Professor Christine Clayton, ZMBH, Heidelberg, for comments and careful reading of the manuscript. V. Starkuviene and C. Claas are supported by BMBF FORSYS ViroQuant (#0313923), BMBF SysTec (0315523A), and Stiftung Baden Württemberg (P-LS-Meth/11). A. Serva is supported by LGFG fellowship of Graduate Academy of Heidelberg University and HBIGS. A. Serva and C. Claas contributed equally to this work.

References

[1] A. S. Flynt and E. C. Lai, "Biological principles of microRNA-mediated regulation: shared themes amid diversity," *Nature Reviews Genetics*, vol. 9, no. 11, pp. 831–842, 2008.

[2] R. W. Carthew and E. J. Sontheimer, "Origins and mechanisms of miRNAs and siRNAs," *Cell*, vol. 136, no. 4, pp. 642–655, 2009.

[3] R. C. Friedman, K. K. Farh, C. B. Burge, and D. P. Bartel, "Most mammalian mRNAs are conserved targets of microRNAs," *Genome Research*, vol. 19, no. 1, pp. 92–105, 2009.

[4] B. P. Lewis, C. B. Burge, and D. P. Bartel, "Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets," *Cell*, vol. 120, no. 1, pp. 15–20, 2005.

[5] J. Brennecke, D. R. Hipfner, A. Stark, R. B. Russell, and S. M. Cohen, "*bantam* encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene *hid in Drosophila*," *Cell*, vol. 113, no. 1, pp. 25–36, 2003.

[6] M. Hackl, S. Brunner, K. Fortschegger et al., "miR-17, miR-19b, miR-20a, and miR-106a are down-regulated in human aging," *Aging Cell*, vol. 9, no. 2, pp. 291–296, 2010.

[7] X. Huang, L. Ding, K. L. Bennewith et al., "Hypoxia-inducible mir-210 regulates normoxic gene expression involved in tumor initiation," *Molecular Cell*, vol. 35, no. 6, pp. 856–867, 2009.

[8] C. D. Johnson, A. Esquela-Kerscher, G. Stefani et al., "The let-7 microRNA represses cell proliferation pathways in human cells," *Cancer Research*, vol. 67, no. 16, pp. 7713–7722, 2007.

[9] Z. Yu, C. Wang, M. Wang et al., "A cyclin D1/microRNA 17/20 regulatory feedback loop in control of breast cancer cell proliferation," *Journal of Cell Biology*, vol. 182, no. 3, pp. 509–517, 2008.

[10] W. Hu, C. S. Chan, R. Wu et al., "Negative regulation of tumor suppressor p53 by microRNA miR-504," *Molecular Cell*, vol. 38, no. 5, pp. 689–699, 2010.

[11] H. Nakano, T. Miyazawa, K. Kinoshita, Y. Yamada, and T. Yoshida, "Functional screening identifies a microRNA, miR-491 that induces apoptosis by targeting Bcl-X(L)in colorectal cancer cells," *International Journal of Cancer*, vol. 127, no. 5, pp. 1072–1080, 2010.

[12] A. K. Pandey, P. Agarwal, K. Kaur, and M. Datta, "MicroRNAs in diabetes: tiny players in big disease," *Cellular Physiology and Biochemistry*, vol. 23, no. 4–6, pp. 221–232, 2009.

[13] H. Y. Ling, H. S. Ou, S. D. Feng et al., "Changes in microrna (mir) profile and effects of mir-320 in insulin-resistant 3t3-l1 adipocytes," *Clinical and Experimental Pharmacology and Physiology*, vol. 36, no. 9, pp. e32–e39, 2009.

[14] X. Tang, G. Tang, and S. Özcan, "Role of microRNAs in diabetes," *Biochimica et Biophysica Acta*, vol. 1779, no. 11, pp. 697–701, 2008.

[15] C. M. Croce, "Causes and consequences of microRNA dysregulation in cancer," *Nature Reviews Genetics*, vol. 10, no. 10, pp. 704–714, 2009.

[16] S. S. Myatt, J. Wang, L. J. Monteiro et al., "Definition of microRNAs that repress expression of the tumor suppressor gene *FOXO1* in endometrial cancer," *Cancer Research*, vol. 70, no. 1, pp. 367–377, 2010.

[17] C. L. Jopling, M. Yi, A. M. Lancaster, S. M. Lemon, and P. Sarnow, "Molecular biology: modulation of hepatitis C virus RNA abundance by a liver-specific microRNA," *Science*, vol. 309, no. 5740, pp. 1577–1581, 2005.

[18] S. Chang, S. Wen, D. Chen, and P. Jin, "Small regulatory RNAs in neurodevelopmental disorders," *Human Molecular Genetics*, vol. 18, no. 1, pp. R18–R26, 2009.

[19] Y. Xu, F. Li, B. Zhang et al., "MicroRNAs and target site screening reveals a pre-microRNA-30e variant associated with schizophrenia," *Schizophrenia Research*, vol. 119, no. 1–3, pp. 219–227, 2010.

[20] F. Takeshita, L. Patrawala, M. Osaki et al., "Systemic delivery of synthetic microRNA-16 inhibits the growth of metastatic prostate tumors via downregulation of multiple cell-cycle genes," *Molecular Therapy*, vol. 18, no. 1, pp. 181–187, 2010.

[21] J. Kota, R. R. Chivukula, K. A. O'Donnell et al., "Therapeutic microRNA delivery suppresses tumorigenesis in a murine liver cancer model," *Cell*, vol. 137, no. 6, pp. 1005–1017, 2009.

[22] S. Volinia, G. A. Calin, C. G. Liu et al., "A microRNA expression signature of human solid tumors defines cancer gene targets," *Proceedings of the National Academy of Sciences*

of the United States of America, vol. 103, no. 7, pp. 2257–2261, 2006.

[23] S. Toffanin, Y. Hoshida, A. Lachenmayer et al., "MicroRNA-based classification of hepatocellular carcinoma and oncogenic role of miR-517a," *Gastroenterology*, vol. 140, no. 5, pp. 1618–1628.e16, 2011.

[24] E. Huntzinger and E. Izaurralde, "Gene silencing by microRNAs: contributions of translational repression and mRNA decay," *Nature Reviews Genetics*, vol. 12, no. 2, pp. 99–110, 2011.

[25] J. Krol, I. Loedige, and W. Filipowicz, "The widespread regulation of microRNA biogenesis, function and decay," *Nature Reviews Genetics*, vol. 11, no. 9, pp. 597–610, 2010.

[26] A. M. Denli, B. B. Tops, R. H. Plasterk, R. F. Ketting, and G. J. Hannon, "Processing of primary microRNAs by the microprocessor complex," *Nature*, vol. 432, no. 7014, pp. 231–235, 2004.

[27] R. Yi, Y. Qin, I. G. Macara, and B. R. Cullen, "Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs," *Genes and Development*, vol. 17, no. 24, pp. 3011–3016, 2003.

[28] Y. Lee, I. Hur, S.-Y. Park, Y.-K. Kim, M. R. Suh, and V. N. Kim, "The role of PACT in the RNA silencing pathway," *EMBO Journal*, vol. 25, no. 3, pp. 522–532, 2006.

[29] T. P. Chendrimada, R. I. Gregory, E. Kumaraswamy et al., "TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing," *Nature*, vol. 436, no. 7051, pp. 740–744, 2005.

[30] D. W. Salzman, J. Shubert-Coleman, and H. Furneaux, "P68 RNA helicase unwinds the human let-7 microRNA precursor duplex and is required for let-7-directed silencing of gene expression," *The Journal of Biological Chemistry*, vol. 282, no. 45, pp. 32773–32779, 2007.

[31] G. B. Robb and T. M. Rana, "RNA helicase a interacts with RISC in human cells and functions in RISC loading," *Molecular Cell*, vol. 26, no. 4, pp. 523–537, 2007.

[32] J. S. Yang, M. D. Phillips, D. Betel et al., "Widespread regulatory activity of vertebrate microRNA* species," *RNA*, vol. 17, no. 2, pp. 312–326, 2011.

[33] G. Meister, M. Landthaler, A. Patkaniowska, Y. Dorsett, G. Teng, and T. Tuschl, "Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs," *Molecular Cell*, vol. 15, no. 2, pp. 185–197, 2004.

[34] A. Eulalio, E. Huntzinger, and E. Izaurralde, "GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay," *Nature Structural and Molecular Biology*, vol. 15, no. 4, pp. 346–353, 2008.

[35] J. Winter, S. Jung, S. Keller, R. I. Gregory, and S. Diederichs, "Many roads to maturity: microRNA biogenesis pathways and their regulation," *Nature Cell Biology*, vol. 11, no. 3, pp. 228–234, 2009.

[36] S. Cheloufi, C. O. Dos Santos, M. M. Chong, and G. J. Hannon, "A dicer-independent miRNA biogenesis pathway that requires Ago catalysis," *Nature*, vol. 465, no. 7298, pp. 584–589, 2010.

[37] D. Cifuentes, H. Xue, D. W. Taylor et al., "A novel miRNA processing pathway independent of dicer requires Argonaute2 catalytic activity," *Science*, vol. 328, no. 5986, pp. 1694–1698, 2010.

[38] S. Y. Park, J. H. Lee, M. Ha, J. W. Nam, and V. N. Kim, "miR-29 miRNAs activate p53 by targeting p85α and CDC42," *Nature Structural and Molecular Biology*, vol. 16, no. 1, pp. 23–29, 2009.

[39] S. Tian, S. Huang, S. Wu, W. Guo, J. Li, and X. He, "MicroRNA-1285 inhibits the expression of p53 by directly targeting its 3' untranslated region," *Biochemical and Biophysical Research Communications*, vol. 396, no. 2, pp. 435–439, 2010.

[40] P. M. Voorhoeve, C. le Sage, M. Schrier et al., "A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors," *Cell*, vol. 124, no. 6, pp. 1169–1181, 2006.

[41] D. Baek, J. Villén, C. Shin, F. D. Camargo, S. P. Gygi, and D. P. Bartel, "The impact of microRNAs on protein output," *Nature*, vol. 455, no. 7209, pp. 64–71, 2008.

[42] M. Selbach, B. Schwanhäusser, N. Thierfelder, Z. Fang, R. Khanin, and N. Rajewsky, "Widespread changes in protein synthesis induced by microRNAs," *Nature*, vol. 455, no. 7209, pp. 58–63, 2008.

[43] N. Bushati and S. M. Cohen, "MicroRNA functions," *Annual Review of Cell and Developmental Biology*, vol. 23, pp. 175–205, 2007.

[44] L. Du, J. J. Schageman, M. C. Subauste et al., "MiR-93, miR-98, and miR-197 regulate expression of tumor suppressor gene *FUS1*," *Molecular Cancer Research*, vol. 7, no. 8, pp. 1234–1243, 2009.

[45] C. le Sage, R. Nagel, D. A. Egan et al., "Regulation of the $p27^{Kip1}$ tumor suppressor by miR-221 and miR-222 promotes cancer cell proliferation," *The EMBO Journal*, vol. 26, no. 15, pp. 3699–3708, 2007.

[46] S. Wu, S. Huang, J. Ding et al., "Multiple microRNAs modulate p21Cip1/Waf1 expression by directly targeting its 3′ untranslated region," *Oncogene*, vol. 29, no. 15, pp. 2302–2308, 2010.

[47] R. S. Pillai, "MicroRNA function: multiple mechanisms for a tiny RNA?" *RNA*, vol. 11, no. 12, pp. 1753–1761, 2005.

[48] I. Lee, S. S. Ajay, J. I. Yook et al., "New class of microRNA targets containing simultaneous 5′-UTR and 3′-UTR interaction sites," *Genome Research*, vol. 19, no. 7, pp. 1175–1183, 2009.

[49] I. Elcheva, S. Goswami, F. K. Noubissi, and V. S. Spiegelman, "CRD-BP protects the coding region of βTrCP1 mRNA from miR-183-mediated degradation," *Molecular Cell*, vol. 35, no. 2, pp. 240–246, 2009.

[50] R. C. Lee, R. L. Feinbaum, and V. Ambros, "The C. elegans heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*," *Cell*, vol. 75, no. 5, pp. 843–854, 1993.

[51] A. Fire, S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello, "Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*," *Nature*, vol. 391, no. 6669, pp. 806–811, 1998.

[52] G. Meister and T. Tuschl, "Mechanisms of gene silencing by double-stranded RNA," *Nature*, vol. 431, no. 7006, pp. 343–349, 2004.

[53] C. C. Mello and D. Conte Jr., "Revealing the world of RNA interference," *Nature*, vol. 431, no. 7006, pp. 338–342, 2004.

[54] Y. Tomari and P. D. Zamore, "Perspective: machines for RNAi," *Genes and Development*, vol. 19, no. 5, pp. 517–529, 2005.

[55] S. M. Hammond, E. Bernstein, D. Beach, and G. J. Hannon, "An RNA-directed nuclease mediates post-transcriptional gene silencing in cells," *Nature*, vol. 404, no. 6775, pp. 293–296, 2000.

[56] Y. Zeng, R. Yi, and B. R. Cullen, "MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar

mechanisms," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 17, pp. 9779–9784, 2003.

[57] W. Filipowicz, S. N. Bhattacharyya, and N. Sonenberg, "Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?" *Nature Reviews Genetics*, vol. 9, no. 2, pp. 102–114, 2008.

[58] C. Collinet, M. Stöter, C. R. Bradshaw et al., "Systems survey of endocytosis by multiparametric image analysis," *Nature*, vol. 464, no. 7286, pp. 243–249, 2010.

[59] B. Neumann, T. Walter, J. K. Hériché et al., "Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes," *Nature*, vol. 464, no. 7289, pp. 721–727, 2010.

[60] M. Boutros and J. Ahringer, "The art and design of genetic screens: RNA interference," *Nature Reviews Genetics*, vol. 9, no. 7, pp. 554–566, 2008.

[61] L. Timmons, D. L. Court, and A. Fire, "Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*," *Gene*, vol. 263, no. 1-2, pp. 103–112, 2001.

[62] L. Timmons and A. Fire, "Specific interference by ingested dsRNA," *Nature*, vol. 395, no. 6705, p. 854, 1998.

[63] J. C. Clemens, C. A. Worby, N. Simonson-Leff et al., "Use of double-stranded RNA interference in *Drosophila* cell lines to dissect signal transduction pathways," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 12, pp. 6499–6503, 2000.

[64] S. M. Elbashir, J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, and T. Tuschl, "Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells," *Nature*, vol. 411, no. 6836, pp. 494–498, 2001.

[65] R. Kittler, L. Pelletier, A. K. Heninger et al., "Genome-scale RNAi profiling of cell division in human tissue culture cells," *Nature Cell Biology*, vol. 9, no. 12, pp. 1401–1412, 2007.

[66] R. König, C. Y. Chiang, B. P. Tu et al., "A probability-based approach for the analysis of large-scale RNAi screens," *Nature Methods*, vol. 4, no. 10, pp. 847–849, 2007.

[67] P. J. Paddison, J. M. Silva, D. S. Conklin et al., "A resource for large-scale RNA-interference-based screens in mammals," *Nature*, vol. 428, no. 6981, pp. 427–431, 2004.

[68] K. Berns, E. M. Hijmans, J. Mullenders et al., "A large-scale RNAi screen in human cells identifies new components of the p53 pathway," *Nature*, vol. 428, no. 6981, pp. 431–437, 2004.

[69] J. Moffat, D. A. Grueneberg, X. Yang et al., "A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen," *Cell*, vol. 124, no. 6, pp. 1283–1298, 2006.

[70] J. M. Silva, M. Z. Li, K. Chang et al., "Second-generation shRNA libraries covering the mouse and human genomes," *Nature Genetics*, vol. 37, no. 11, pp. 1281–1288, 2005.

[71] M. Garofalo and C. M. Croce, "MicroRNAs: master regulators as potential therapeutics in cancer," *Annual Review of Pharmacology and Toxicology*, vol. 51, pp. 25–43, 2011.

[72] K. J. Mavrakis, A. L. Wolfe, E. Oricchio et al., "Genome-wide RNA-mediated interference screen identifies miR-19 targets in notch-induced T-cell acute lymphoblastic leukaemia," *Nature Cell Biology*, vol. 12, no. 4, pp. 372–379, 2010.

[73] D. P. Bartel, "MicroRNAs: target recognition and regulatory functions," *Cell*, vol. 136, no. 2, pp. 215–233, 2009.

[74] A. Krek, D. Grün, M. N. Poy et al., "Combinatorial microRNA target predictions," *Nature Genetics*, vol. 37, no. 5, pp. 495–500, 2005.

[75] B. John, A. J. Enright, A. Aravin, T. Tuschl, C. Sander, and D. S. Marks, "Human microRNA targets," *PLoS Biology*, vol. 2, no. 11, article e363, 2004.

[76] X. Xie, J. Lu, E. J. Kulbokas et al., "Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals," *Nature*, vol. 434, no. 7031, pp. 338–345, 2005.

[77] M. R. Capecchi, "Generating mice with targeted mutations," *Nature Medicine*, vol. 7, no. 10, pp. 1086–1090, 2001.

[78] C. Tong, P. Li, N. L. Wu, Y. Yan, and Q. L. Ying, "Production of p53 gene knockout rats by homologous recombination in embryonic stem cells," *Nature*, vol. 467, no. 7312, pp. 211–213, 2010.

[79] F. D. Urnov, E. J. Rebar, M. C. Holmes, H. S. Zhang, and P. D. Gregory, "Genome editing with engineered zinc finger nucleases," *Nature Reviews Genetics*, vol. 11, no. 9, pp. 636–646, 2010.

[80] P. C. Zamecnik and M. L. Stephenson, "Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 75, no. 1, pp. 280–284, 1978.

[81] J. Kurreck, "Antisense technologies: improvement through novel chemical modifications," *European Journal of Biochemistry*, vol. 270, no. 8, pp. 1628–1644, 2003.

[82] T. A. Vickers, S. Koo, C. F. Bennett, S. T. Crooke, N. M. Dean, and B. F. Baker, "Efficient reduction of target RNAs by small interfering RNA and RNase H-dependent antisense agents. A comparative analysis," *The Journal of Biological Chemistry*, vol. 278, no. 9, pp. 7108–7118, 2003.

[83] J. Krützfeldt, S. Kuwajima, R. Braich et al., "Specificity, duplex degradation and subcellular localization of antagomirs," *Nucleic Acids Research*, vol. 35, no. 9, pp. 2885–2892, 2007.

[84] J. Elmén, M. Lindow, A. Silahtaroglu et al., "Antagonism of microRNA-122 in mice by systemically administered LNA-antimiR leads to up-regulation of a large set of predicted target mRNAs in the liver," *Nucleic Acids Research*, vol. 36, no. 4, pp. 1153–1162, 2008.

[85] S. Davis, S. Propp, S. M. Freier et al., "Potent inhibition of microRNA in vivo without degradation," *Nucleic Acids Research*, vol. 37, no. 1, pp. 70–77, 2009.

[86] Y. Lu, J. Xiao, H. Lin et al., "A single anti-microRNA antisense oligodeoxyribonucleotide (AMO) targeting multiple microRNAs offers an improved approach for microRNA interference," *Nucleic Acids Research*, vol. 37, no. 3, article e24, 2009.

[87] C. C. Esau, "Inhibition of microRNA with antisense oligonucleotides," *Methods*, vol. 44, no. 1, pp. 55–60, 2008.

[88] W. P. Kloosterman, A. K. Lagendijk, R. F. Ketting, J. D. Moulton, and R. H. Plasterk, "Targeted inhibition of miRNA maturation with morpholinos reveals a role for miR-375 in pancreatic islet development," *PLoS Biology*, vol. 5, no. 8, article e203, 2007.

[89] Y. S. Lee, H. K. Kim, S. Chung, K. S. Kim, and A. Dutta, "Depletion of human micro-RNA miR-125b reveals that it is critical for the proliferation of differentiated cells but not for the down-regulation of putative targets during differentiation," *The Journal of Biological Chemistry*, vol. 280, no. 17, pp. 16635–16641, 2005.

[90] J. S. Eisen and J. C. Smith, "Controlling morpholino experiments: don't stop making antisense," *Development*, vol. 135, no. 10, pp. 1735–1743, 2008.

[91] M. S. Ebert, J. R. Neilson, and P. A. Sharp, "MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells," *Nature Methods*, vol. 4, no. 9, pp. 721–726, 2007.

[92] C. M. Loya, C. S. Lu, D. Van Vactor, and T. A. Fulga, "Transgenic microRNA inhibition with spatiotemporal specificity in intact organisms," *Nature Methods*, vol. 6, no. 12, pp. 897–903, 2009.

[93] N. Tsuda, S. Ishiyama, Y. Li, C. G. Ioannides, J. L. Abbruzzese, and D. Z. Chang, "Synthetic microRNA designed to target glioma-associated antigen 1 transcription factor inhibits division and induces late apoptosis in pancreatic tumor cells," *Clinical Cancer Research*, vol. 12, no. 21, pp. 6557–6564, 2006.

[94] R. Garzon, C. E. Heaphy, V. Havelange et al., "MicroRNA 29b functions in acute myeloid leukemia," *Blood*, vol. 114, no. 26, pp. 5331–5341, 2009.

[95] K. H. Chung, C. C. Hart, S. Al-Bassam et al., "Polycistronic RNA polymerase II expression vectors for RNA interference based on BIC/miR-155," *Nucleic Acids Research*, vol. 34, no. 7, p. e53, 2006.

[96] D. Kumar, C. I. An, and Y. Yokobayashi, "Conditional RNA interference mediated by allosteric ribozyme," *Journal of the American Chemical Society*, vol. 131, no. 39, pp. 13906–13907, 2009.

[97] M. Schmidt-Supprian and K. Rajewsky, "Vagaries of conditional gene targeting," *Nature Immunology*, vol. 8, no. 7, pp. 665–668, 2007.

[98] W. Y. Choi, A. J. Giraldez, and A. F. Schier, "Target protectors reveal dampening and balancing of nodal agonist and antagonist by miR-430," *Science*, vol. 318, no. 5848, pp. 271–274, 2007.

[99] L. T. Lam, X. Lu, H. Zhang, R. R. Lesniewski, S. H. Rosenberg, and D. Semizarov, "A microRNA screen to identify modulators of sensitivity to BCL2 inhibitor ABT-263 (navitoclax)," *Molecular Cancer Therapeutics*, vol. 9, no. 11, pp. 2943–2950, 2010.

[100] A. M. Cheng, M. W. Byrom, J. Shelton, and L. P. Ford, "Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis," *Nucleic Acids Research*, vol. 33, no. 4, pp. 1290–1297, 2005.

[101] Q. Huang, K. Gumireddy, M. Schrier et al., "The microRNAs miR-373 and miR-520c promote tumour invasion and metastasis," *Nature Cell Biology*, vol. 10, no. 2, pp. 202–210, 2008.

[102] R. Nagel, L. Clijsters, and R. Agami, "The miRNA-192/194 cluster regulates the *Period* gene family and the circadian clock," *FEBS Journal*, vol. 276, no. 19, pp. 5447–5455, 2009.

[103] T. Subramanian and G. Chinnadurai, "Temperature-sensitive replication-competent adenovirus shRNA vectors to study cellular genes in virus-induced apoptosis," *Methods in Molecular Medicine*, vol. 130, pp. 125–134, 2007.

[104] D. Ovcharenko, K. Kelnar, C. Johnson, N. Leng, and D. Brown, "Genome-scale microRNA and small interfering RNA screens identify small RNA modulators of TRAIL-induced apoptosis pathway," *Cancer Research*, vol. 67, no. 22, pp. 10782–10788, 2007.

[105] R. Whittaker, P. A. Loy, E. Sisman et al., "Identification of microRNAs that control lipid droplet formation and growth in hepatocytes via high-content screening," *Journal of Biomolecular Screening*, vol. 15, no. 7, pp. 798–805, 2010.

[106] B. Wightman, I. Ha, and G. Ruvkun, "Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in C. elegans," *Cell*, vol. 75, no. 5, pp. 855–862, 1993.

[107] B. J. Reinhart, F. J. Slack, M. Basson et al., "The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*," *Nature*, vol. 403, no. 6772, pp. 901–906, 2000.

[108] A. E. Pasquinelli, B. J. Reinhart, F. Slack et al., "Conservation of the sequence and temporal expression of *let-7* heterochronic regulatory RNA," *Nature*, vol. 408, no. 6808, pp. 86–89, 2000.

[109] P. Aza-Blanc, C. L. Cooper, K. Wagner, S. Batalov, Q. L. Deveraux, and M. P. Cooke, "Identification of modulators of TRAIL-induced apoptosis via RNAi-based phenotypic screening," *Molecular Cell*, vol. 12, no. 3, pp. 627–637, 2003.

[110] S. J. Silver, J. W. Hagen, K. Okamura, N. Perrimon, and E. C. Lai, "Functional screening identifies miR-315 as a potent activator of Wingless signaling," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 46, pp. 18151–18156, 2007.

[111] K. A. Cole, E. F. Attiyeh, Y. P. Mosse et al., "A functional screen identifies miR-34a as a candidate neuroblastoma tumor suppressor gene," *Molecular Cancer Research*, vol. 6, no. 5, pp. 735–742, 2008.

[112] A. V. Sirotkin, D. Ovcharenko, R. Grossmann, M. Lauková, and M. Mlynček, "Identification of microRNAs controlling human ovarian cell steroidogenesis via a genome-scale screen," *Journal of Cellular Physiology*, vol. 219, no. 2, pp. 415–420, 2009.

[113] A. V. Sirotkin, M. Lauková, D. Ovcharenko, P. Brenaut, and M. Mlynček, "Identification of microRNAs controlling human ovarian cell proliferation and apoptosis," *Journal of Cellular Physiology*, vol. 223, no. 1, pp. 49–56, 2010.

[114] V. Borgdorff, M. E. Lleonart, C. L. Bishop et al., "Multiple microRNAs rescue from Ras-induced senescence by inhibiting $p21^{Waf1/Cip1}$," *Oncogene*, vol. 29, no. 15, pp. 2262–2271, 2010.

[115] M. Izumiya, K. Okamoto, N. Tsuchiya, and H. Nakagama, "Functional screening using a microRNA virus library and microarrays: a new high-throughput assay to identify tumor-suppressive microRNAs," *Carcinogenesis*, vol. 31, no. 8, pp. 1354–1359, 2010.

[116] A. Serva, B. Knapp, C. Claas et al., "MiR-17 family regulates cargo trafficking," *in review*.

[117] Y. Sylvestre, V. De Guire, E. Querido et al., "An E2F/miR-20a autoregulatory feedback loop," *The Journal of Biological Chemistry*, vol. 282, no. 4, pp. 2135–2143, 2007.

[118] S. Galardi, N. Mercatelli, E. Giorda et al., "miR-221 and miR-222 expression affects the proliferation potential of human prostate carcinoma cell lines by targeting $p27^{Kip1}$," *The Journal of Biological Chemistry*, vol. 282, no. 32, pp. 23716–23724, 2007.

[119] Y. Akao, Y. Nakagawa, and T. Naoe, "*let-7* functions as a potential growth suppressor in human colon cancer cells," *Biological and Pharmaceutical Bulletin*, vol. 29, no. 5, pp. 903–906, 2006.

[120] N. Felli, L. Fontana, E. Pelosi et al., "MicroRNAs 221 and 222 inhibit normal erythropoiesis and erythroleukemic cell growth via kit receptor down-modulation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 50, pp. 18081–18086, 2005.

[121] C. Esau, X. Kang, E. Peralta et al., "MicroRNA-143 regulates adipocyte differentiation," *The Journal of Biological Chemistry*, vol. 279, no. 50, pp. 52361–52365, 2004.

[122] J. A. Chan, A. M. Krichevsky, and K. S. Kosik, "MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells," *Cancer Research*, vol. 65, no. 14, pp. 6029–6033, 2005.

[123] J. F. Chen, E. M. Mandel, J. M. Thomson et al., "The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation," *Nature Genetics*, vol. 38, no. 2, pp. 228–233, 2006.

[124] Y. Hayashita, H. Osada, Y. Tatematsu et al., "A polycistronic microRNA cluster, *miR-17-92*, is overexpressed in human lung cancers and enhances cell proliferation," *Cancer Research*, vol. 65, no. 21, pp. 9628–9632, 2005.

[125] Y. Lu, J. M. Thomson, H. Y. Wong, S. M. Hammond, and B. L. Hogan, "Transgenic over-expression of the microRNA *miR-17-92* cluster promotes proliferation and inhibits differentiation of lung epithelial progenitor cells," *Developmental Biology*, vol. 310, no. 2, pp. 442–453, 2007.

[126] I. Manni, S. Artuso, S. Careccia et al., "The microRNA miR-92 increases proliferation of myeloid cells and by targeting p63 modulates the abundance of its isoforms," *FASEB Journal*, vol. 23, no. 11, pp. 3957–3966, 2009.

[127] M. T. Pickering, B. M. Stadler, and T. F. Kowalik, "MiR-17 and miR-20a temper an E2F1-induced G1 checkpoint to regulate cell cycle progression," *Oncogene*, vol. 28, no. 1, pp. 140–145, 2009.

[128] K. A. O'Donnell, E. A. Wentzel, K. I. Zeller, C. V. Dang, and J. T. Mendell, "c-Myc-regulated microRNAs modulate E2F1 expression," *Nature*, vol. 435, no. 7043, pp. 839–843, 2005.

[129] M. K. Manion and D. M. Hockenbery, "Targeting BCL-2-related proteins in cancer therapy," *Cancer Biology Therapy*, vol. 2, no. 4, pp. S105–S114, 2003.

[130] A. Esquela-Kerscher, P. Trang, J. F. Wiggins et al., "The let-7 microRNA reduces tumor growth in mouse models of lung cancer," *Cell Cycle*, vol. 7, no. 6, pp. 759–764, 2008.

[131] L. Jiang, Q. Huang, S. Zhang et al., "Hsa-miR-125a-3p and hsa-miR-125a-5p are downregulated in non-small cell lung cancer and have inverse effects on invasion and migration of lung cancer cells," *BMC Cancer*, vol. 10, article 318, 2010.

[132] G. Meister, M. Landthaler, Y. Dorsett, and T. Tuschl, "Sequence-specific inhibition of microRNA-and siRNA-induced RNA silencing," *RNA*, vol. 10, no. 3, pp. 544–550, 2004.

[133] R. Pepperkok and J. Ellenberg, "High-throughput fluorescence microscopy for systems biology," *Nature Reviews Molecular Cell Biology*, vol. 7, no. 9, pp. 690–696, 2006.

[134] R. Sacher, L. Stergiou, and L. Pelkmans, "Lessons from genetics: interpreting complex phenotypes in RNAi screens," *Current Opinion in Cell Biology*, vol. 20, no. 4, pp. 483–489, 2008.

[135] N. Rajewsky, "MicroRNA target predictions in animals," *Nature Genetics*, vol. 38, no. 1, supplement, pp. S8–S13, 2006.

[136] M. M. Martin, E. J. Lee, J. A. Buckenberger, T. D. Schmittgen, and T. S. Elton, "MicroRNA-155 regulates human angiotensin II type 1 receptor expression in fibroblast," *The Journal of Biological Chemistry*, vol. 281, no. 27, pp. 18277–18284, 2006.

[137] P. Brodersen and O. Voinnet, "Revisiting the principles of microRNA target recognition and mode of action," *Nature Reviews Molecular Cell Biology*, vol. 10, no. 2, pp. 141–148, 2009.

[138] M. Beitzinger, L. Peters, J. Y. Zhu, E. Kremmer, and G. Meister, "Identification of human microRNA targets from isolated argonaute protein complexes," *RNA Biology*, vol. 4, no. 2, pp. 76–84, 2007.

[139] G. Easow, A. A. Teleman, and S. M. Cohen, "Isolation of microRNA targets by miRNP immunopurification," *RNA*, vol. 13, no. 8, pp. 1198–1204, 2007.

[140] F. V. Karginov, C. Conaco, Z. Xuan et al., "A biochemical approach to identifying microRNA targets," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 49, pp. 19291–19296, 2007.

[141] A. G. Fraser, R. S. Kamath, P. Zipperlen, M. Martinez-Campos, M. Sohrmann, and J. Ahringer, "Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference," *Nature*, vol. 408, no. 6810, pp. 325–330, 2000.

[142] R. S. Kamath, A. G. Fraser, Y. Dong et al., "Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi," *Nature*, vol. 421, no. 6920, pp. 231–237, 2003.

[143] H. Zhou, M. Xu, Q. Huang et al., "Genome-scale RNAi screen for host factors required for HIV replication," *Cell Host and Microbe*, vol. 4, no. 5, pp. 495–504, 2008.

[144] A. E. Carpenter and D. M. Sabatini, "Systematic genome-wide screens of gene function," *Nature Reviews Genetics*, vol. 5, no. 1, pp. 11–22, 2004.

[145] C. Falschlehner, S. Steinbrink, G. Erdmann, and M. Boutros, "High-throughput RNAi screening to dissect cellular pathways: a how-to guide," *Biotechnology Journal*, vol. 5, no. 4, pp. 368–376, 2010.

[146] T. Horn, T. Sandmann, and M. Boutros, "Design and evaluation of genome-wide libraries for RNA interference screens," *Genome Biology*, vol. 11, no. 6, article R61, 2010.

[147] J. T. Marques and B. R. Williams, "Activation of the mammalian immune system by siRNAs," *Nature Biotechnology*, vol. 23, no. 11, pp. 1399–1405, 2005.

[148] A. D. Judge, V. Sood, J. R. Shaw, D. Fang, K. McClintock, and I. MacLachlan, "Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA," *Nature Biotechnology*, vol. 23, no. 4, pp. 457–462, 2005.

[149] J. Elmén, H. Thonberg, K. Ljungberg et al., "Locked nucleic acid (LNA) mediated improvements in siRNA stability and functionality," *Nucleic Acids Research*, vol. 33, no. 1, pp. 439–447, 2005.

[150] A. L. Jackson, J. Burchard, J. Schelter et al., "Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity," *RNA*, vol. 12, no. 7, pp. 1179–1187, 2006.

[151] C. J. Echeverri, P. A. Beachy, B. Baum et al., "Minimizing the risk of reporting false positives in large-scale RNAi screens," *Nature Methods*, vol. 3, no. 10, pp. 777–779, 2006.

[152] C. J. Echeverri and N. Perrimon, "High-throughput RNAi screening in cultured cells: a user's guide," *Nature Reviews Genetics*, vol. 7, no. 5, pp. 373–384, 2006.

[153] D. Grimm, K. L. Streetz, C. L. Jopling et al., "Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways," *Nature*, vol. 441, no. 7092, pp. 537–541, 2006.

[154] A. L. Jackson and P. S. Linsley, "Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application," *Nature Reviews Drug Discovery*, vol. 9, no. 1, pp. 57–67, 2010.

[155] F. Stegmeier, G. Hu, R. J. Rickles, G. J. Hannon, and S. J. Elledge, "A lentiviral microRNA-based system for single-copy polymerase II-regulated RNA interference in mammalian cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 37, pp. 13212–13217, 2005.

[156] D. Grimm and M. A. Kay, "Therapeutic short hairpin RNA expression in the liver: viral targets and vectors," *Gene Therapy*, vol. 13, no. 6, pp. 563–575, 2006.

[157] S. Diederichs, S. Jung, S. M. Rothenberg, G. A. Smolen, B. G. Mlody, and D. A. Haber, "Coexpression of Argonaute-2 enhances RNA interference toward perfect match binding sites," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 27, pp. 9284–9289, 2008.

[158] D. P. Bartel and C. Z. Chen, "Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs," *Nature Reviews Genetics*, vol. 5, no. 5, pp. 396–400, 2004.

[159] E. Del Nery, S. Miserey-Lenkei, T. Falguières et al., "Rab6A and Rab6A′ GTPases play non-overlapping roles in membrane trafficking," *Traffic*, vol. 7, no. 4, pp. 394–407, 2006.

[160] V. Ambros, "MicroRNAs: genetically sensitized worms reveal new secrets," *Current Biology*, vol. 20, no. 14, pp. R598–R600, 2010.

[161] E. A. Miska, E. Alvarez-Saavedra, A. L. Abbott et al., "Most *Caenorhabditis elegans* microRNAs are individually not essential for development or viability," *PLoS genetics*, vol. 3, no. 12, article e215, 2007.

[162] J. L. Brenner, K. L. Jasiewicz, A. F. Fahley, B. J. Kemp, and A. L. Abbott, "Loss of individual microRNAs causes mutant phenotypes in sensitized genetic backgrounds in *C. elegans*," *Current Biology*, vol. 20, no. 14, pp. 1321–1325, 2010.

SAGE-Hindawi Access to Research
Journal of Nucleic Acids
Volume 2011, Article ID 741723, 15 pages
doi:10.4061/2011/741723

Review Article

Molecular Beacons: Powerful Tools for Imaging RNA in Living Cells

Ricardo Monroy-Contreras and Luis Vaca

Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, 04510 Mexico, DF, Mexico

Correspondence should be addressed to Luis Vaca, lvaca@ifc.unam.mx

Received 9 November 2010; Revised 14 June 2011; Accepted 22 June 2011

Academic Editor: Dmitry A. Stetsenko

Recent advances in RNA functional studies highlights the pivotal role of these molecules in cell physiology. Diverse methods have been implemented to measure the expression levels of various RNA species, using either purified RNA or fixed cells. Despite the fact that fixed cells offer the possibility to observe the spatial distribution of RNA, assays with capability to real-time monitoring RNA transport into living cells are needed to further understand the role of RNA dynamics in cellular functions. Molecular beacons (MBs) are stem-loop hairpin-structured oligonucleotides equipped with a fluorescence quencher at one end and a fluorescent dye (also called reporter or fluorophore) at the opposite end. This structure permits that MB in the absence of their target complementary sequence do not fluoresce. Upon binding to targets, MBs emit fluorescence, due to the spatial separation of the quencher and the reporter. Molecular beacons are promising probes for the development of RNA imaging techniques; nevertheless much work remains to be done in order to obtain a robust technology for imaging various RNA molecules together in real time and in living cells. The present work concentrates on the different requirements needed to use successfully MB for cellular studies, summarizing recent advances in this area.

1. Introduction

A wealth of experimental evidence accumulated to this date illustrates the wide range of cellular functions conducted by ribonucleic acids (RNAs). In many occasions these functions involve RNA transport from one cellular compartment to another or into the same compartment to specialized regions. In order to accomplish this, RNA molecules are under a specific and selective control via expression levels and/or stability in a spatial-temporal manner. Different methods developed to purify RNA [1–3] from cell populations or tissues have provided relevant information about the relative concentration of RNA in cells, or cellular compartments, yet these methods provide limited information about the spatial-temporal distribution of RNAs and their dynamic transport.

The integral understanding of cellular process in which RNA is involved, requires a method that reveals RNA localization in real time in a subcellular context in living cells. The information obtained from this type of assays promise to impulse the advancement in molecular biology, medical research, and diagnostics.

Many methods have been developed to measure RNA expression levels between different cell populations, such as the polymerase chain reaction (PCR) [4], northern blot [5], serial analysis of the gene expression (SAGE) [6], suppression of subtractive hybridization (SSH) [7], differential display [8], representational difference analysis (RDA) [9], expressed sequence tag (EST) [10], and microarrays [11]. Thus these techniques in addition to databases containing genomic data for a wide variety of biological entities, supply a useful instrument for understanding physiology or pathology at the molecular level. Nevertheless, none of the techniques previously mentioned could provide information about the subcellular RNA localization and its transport.

Through the use of fluorescent *in situ* hybridization assays [12], scientists have identified gradients of RNA in cells [13], embryos [14], and tissues [15, 16]. However this technique is conducted on fixed samples and therefore provides limited information about the dynamics of such

gradients. Other disadvantages are the laborious and time-consuming nature of the assay, the difficulty that involves the analysis of the images and variations due to handling different trials. The chemicals used for dehydration and fixation could affect the signal level [17] and alter the integrity of organelles, hampering any conclusions about spatial changes in organelles or compartments.

The most adequate procedure to obtain spatial-temporal resolution of the RNA dynamics is using living cell assays. Under these conditions, the probes to be used must have high specificity, sensitivity, and good signal to background ratio, especially for low abundance RNA molecules. Ideally, the experimental procedure should be able to detect discrete changes in RNA concentrations of the sample. Furthermore, a robust method should provide additional information like single nucleotide mutations, deletions, insertions, and single nucleotide polymorphisms (SNPs). Another fact to take in consideration is the method for intracellular delivery; it should not produce the degradation of the probe or damage the cell. Finally, the probes utilized should have fast kinetics of hybridization with its target at room temperature, in order to be useful in real-time measurements.

Recent advances in DNA chemistry deliver promising nanostructured probes known as molecular beacons (MB), which may fulfill many (if not all) of the requirements highlighted above. MBs are slowly becoming powerful tools to explore RNA function and dynamics in living cells. Combined with advanced imaging techniques, MB are delivering interesting and sometimes surprising results about RNA dynamics in living cells and tissues.

A recent work reviews different forms of linear and nonlinear probes useful for cell imaging studies [18]. In this review we will address the use of MBs for RNA imaging in living cells, analyzing some of the key features required for successful MB design. We will review also various delivery methods commonly used for MB introduction into living cells.

2. Structure and Function of Molecular Beacons

MBs are oligonucleotide probes that fluoresce upon hybridization with its targets [20]. To build a MB, an oligonucleotide chain is linked covalently to a fluorescent dye (reporter) usually at 5′ end and to a fluorescence quencher at the opposite end. The probes have the capability to produce a stem-loop hairpin structural conformation in absence of its target, due to complementary sequences from 5 to 7 nucleotides in length, located at the termini of the molecule (Figure 1(a)). The resulting secondary structure of the MB brings together the reporter and the quencher, making possible the energy transfer, which produces fluorescence quenching. When MB hybridizes to its target, the reporter and quencher are separated ($\geq$10 nm), preventing the quenching of the fluorescence from the reporter (Figure 1(b)).

Most MB studies have been carried out using nucleic acids in *in vitro* assays. In fact, several real-time PCR strategies involve the use of MB for monitoring sample amplification [21–25]. Such strategies involve heating of the solution to high degrees in order to maximize MB-target association. Unfortunately, heating to high temperatures cannot be conducted when studying living cells, and therefore MB design should account for MB-target association at physiological temperature.

We have found that a careful design of the MB stem sequence is essential for MB-target association at physiological temperature ranging from 20 to 37 degrees centigrade.

In a solution of MB containing an excess of target ($\geq$fivefold MB concentration) it is possible to observe an increment in fluorescence intensity between 10- and 200-fold [20] in comparison with a solution without a target (Figure 1(c)). When target concentration is unknown (like in the case of nucleic acids detection in living cells) it is desirable to have fast hybridization rates and improved signal to background ratio, especially when the target is a nucleic acid present in a low concentration. Some MB design strategies have been implemented recently to address this issue.

MB shows an increased specificity in comparison to linear probes having an equal number of complementary nucleotides with its target; this is due to competition between the stem-loop hairpin structure and the MB-target hybridized state [26]. This phenomenon gives the capability to the MB to discriminate between perfectly matched and single mismatched targets. To reach this high specificity an appropriate MB design is required, taking into consideration the experimental conditions, such as magnesium concentration, ionic strength, and temperature.

We have found that prior to living cell studies, it is highly advisable to conduct hybridization and thermal denaturing assays using a spectrofluorometer and a real-time thermal cycler, respectably. In order to have a prediction about the MB behavior in living cells assays it is recommended to calculate the thermodynamical parameters of the hybridization kinetics, that is, changes in enthalpy and entropy, melting temperature, also association and dissociation constants of a specific MB-target pair using the thermal profile obtained from the thermal denaturing and parameters obtained from the hybridization assays. See [26, 27] for detailed protocols.

Molecular beacons have been used successfully in a wide variety of applications due to its high signal to background ratio, including qPCR [21–25], SNPs detection by real-time PCR [28–31], genotyping [32–36], mutation detection [37–39], assaying for nucleic acid cleavage [40, 41], cancer-cell detection [42–44], monitoring viral infection [45–47], and RNA expression, transport, and localization in living cells [48–51].

3. Designing Molecular Beacons

Three key factors must be considered for MB design, the efficiency of the quenching, the stem-loop secondary structure formed by the sequences at the MB termini and the C-G content of the stem (specially relevant for assays conducted at physiological temperatures).

3.1. Reporter and Quencher: Mechanism and Pair Election. MB use a reporter-quencher pair to build a signal transduction system [52, 53]. The quenching that takes place when

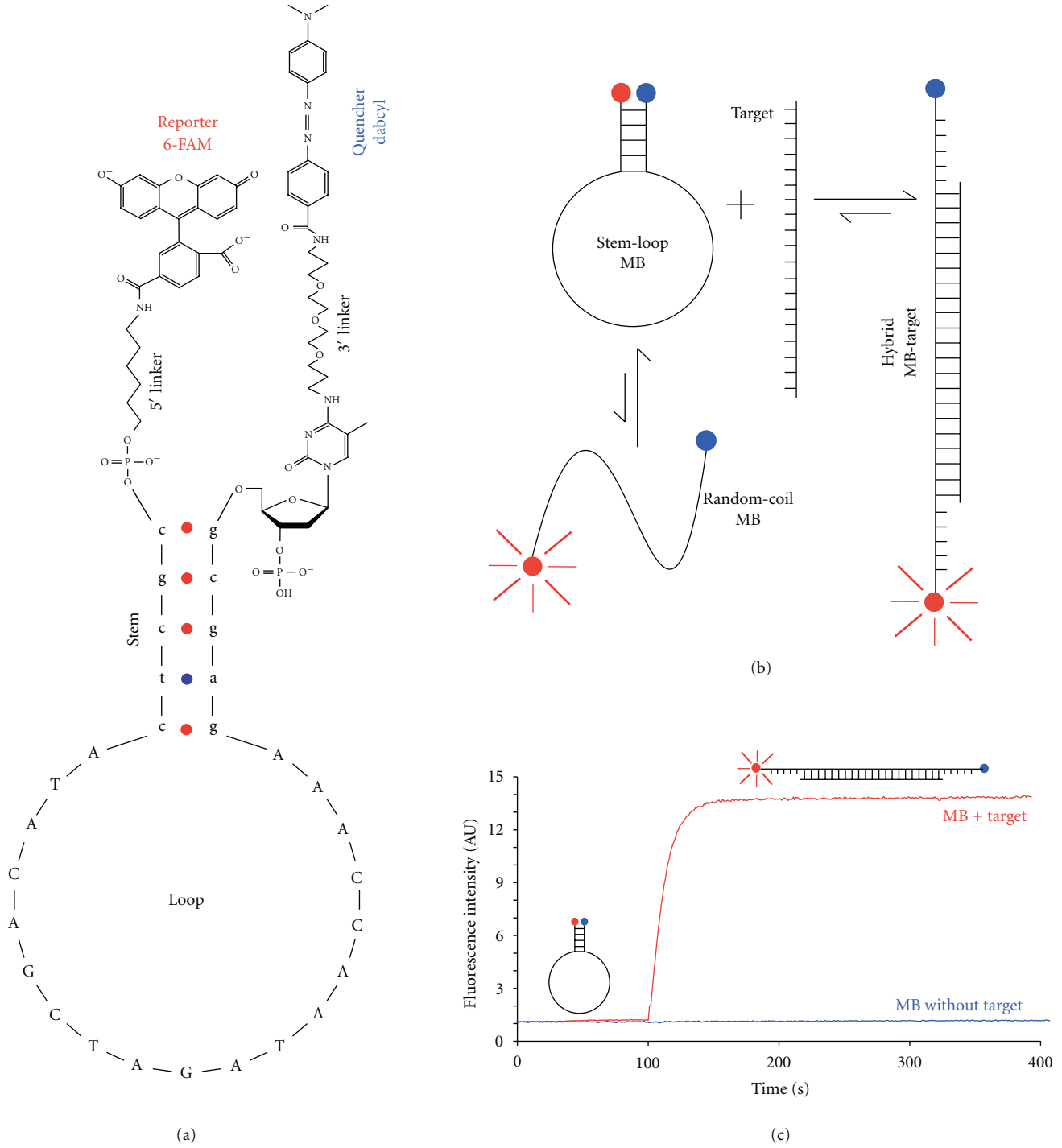


Figure 1: Structure and function of MB. (a) Stem-loop hairpin structure of a MB showing its four structural components: loop, stem, quencher, and reporter. The chemical structure of the linkers is drawn according to the manufacturer (Integrated DNA Technologies, Iowa, USA). (b) Mechanism of MB function. A MB in a solution containing both MB and target could be in three states: free in a stem-loop hairpin conformation, hybridized with it target, or unbound in a random-coil conformation. The random-coil conformation of the MB contributes to the background. (c) Fluorescence intensity. The blue line shows the fluorescence intensity for a MB in solution (50 nM) during 400 seconds (background), and the red line corresponds to the addition (100 seconds) of a target oligonucleotide (500 nM) that hybridize the MB at the loop region. An increase in fluorescent intensity of approximately thirteenfold is observed. The reporter is represented with a red circle and the quencher with a blue. UA means arbitrary units.

Table 1: Reporter-quencher pairs. For every one reporter is showed the peak of excitation, emission, and also the quenchers that could be used for observing a high quenching efficiency. The excitation peak and emission peak values were obtained from reference [58].

Reporter	Excitation peak (nm)	Emission peak (nm)	Compatible quenchers
Cy3	550	605	BHQ-2 and Dabcyl
Cy5	640	675	BHQ-3 and Dabcyl
CR-6G	505	570	Dabcyl
6-FAM	484	525	Iowa Black-FQ, BHQ-1 and Dabcyl
HEX	525	554	BHQ-1, Iowa Black-FQ and Dabcyl
TAMRA	555	605	Dabcyl
TET	490	535	BHQ-1, Iowa Black-FQ and Dabcyl

the MB is in its stem-loop structural conformation (hairpin) occurs by two mechanisms.

The first mechanism is called dynamic quenching and includes Föster's resonance energy transfer (FRET) and Dexter transfer, also named electron transfer quenching. In the dynamic quenching mechanism, the photon from the reporter is not released to the environment because a long-range dipole-dipole interaction occurs between reporter and quencher. For the dynamic quenching mechanism, the transfer efficiency of the energy depends on the spectral overlap between the emission spectrum of the reporter and the absorption spectrum of the quencher [54], the quantum yield of the donor, the relative orientation (also referred to as dipolar moment), and the distance between the two groups [55]. For a 50% of FRET efficiency, a distance in the range of 20–70 Å (Föster's distance) is needed between the reporter and quencher, losing all the quenching efficiency at a distance ≥100 Å (10 nm).

The second mechanism called static quenching [56], requires the formation of ground-state complexes and the mechanism depends on the stem sequence, the linkers, and reporter-quencher pair used in the MB. The most important contributor in the quenching phenomenon is the static or contact quenching; the name describes the nature of the process, which occurs only when reporter and quencher are in close proximity, permitting the physical contact between them. Under these circumstances most of the energy transferred is not emitted as light, but emitted in the form of heat through a nonirradiative process [57]. This phenomenon occurs for all the nonfluorescent quenchers and determines basal fluorescence value in the absence of target and the signal-to-noise ratio of the MB. The most frequently used nonfluorescent quenchers in MB synthesis are Dabcyl (4-([4′-(dimethylamino)phenyl]azo)benzoic acid), Black-Hole quenchers BHQ1 and BHQ2, and Iowa Black FQ and RQ. Typical static quenching efficiency for these quenchers is in the range of 85–97% [54, 57]. These quenchers could be paired with various reporters (Table 1). Dabcyl is practically a universal quencher for near-infrared reporters although its absorption peak is around 475 nm. Worth mentioning, this quencher is the least expensive of all.

Most recently, it has been reported that guanine bases can be utilized as quenchers in MB design (a method known as G-quenched MB) [59, 60]. Guanine bases work as quenchers when present at the opposite end of the reporter. The guanine bases could be placed as overhang or not. MBs without overhanged guanines show only a 15% increment in fluorescence intensity moving from the unbound state (in the absence of target) to the target-hybridized form [61], but when using two overhanged guanines a 2-fold is observed [62]. Guanine bases function as quenchers for a variety of reporters, including pyrene, coumarin, acridone, and 6-FAM. Using guanine bases as quenchers improves the yield of MB obtained during the synthesis and reduces significantly its cost. G-quenched MBs have not been used for RNA imaging in live-cell assays, it is a field waiting to be explored. The structure and function of G-quenched MB and other quencher-free probes are widely discussed in reference [63].

3.2. Thermodynamics and Design. The MB-target hybridization is due to Watson-Crick basepairing. Experimental evidence (thermal denaturing assays) and binary models that simplify the real behavior of the phenomenon considering only the beginning of the hybridization kinetics and the end (steady-state), support the fact that for a given MB-perfect-matched target pair the melting temperature is 13 ± 3°C higher than for the corresponding single-mismatched target [26, 27, 57]. The capability of a probe to discriminate perfect a perfect matched and single mismatched target is known as selectivity potential, and its magnitude is determined by the difference of the temperatures at the transition phases [64].

Linear probes have just two possible states, one free or unbound and the other bound to a target. Whereas MBs have three possible states, free with stem-loop conformation, free as a random coil and bound to its target (Figure 1(b)). In terms of free energy a MB is most stable in stem-loop conformation (hairpin) than in random coil state; therefore, such transitions most be brief. These suggest that when the temperature increases and the dissociation between the target and the MB occurs, the MB will acquire the stem loop conformation preferentially over the random coil state [65–67]. The previously described phenomenon explains the high signal to background ratio observed with MBs. This suggest also that a more stable stem (due to higher C-G content or stem length) will produce a higher signal to background ratio. MBs with more stable stems would show enhanced

selectivity because the MB will hybridize only when the interactions MB-targets are strong enough to overcome the stem stability. However, a word of caution, MBs with more stable stems may not hybridize to its target at a physiological temperature, preventing its use for cellular studies. Thus a delicate balance between hairpin stability (stem design) and MB-target selectivity must be reached for successful use of MBs in living cell assays.

3.3. Balance between Selectivity and Hybridization Rate. Clearly the greater advantage of MB over linear probes is its enhanced selectivity, which result from the hairpin stem-loop structure attained by the MB in the nonhybridized state (not associated to its target). The stability of the hairpin conformation is ensured by building more stable stems (either increasing its C-G content or the stem length). Unfortunately, increasing stem strength decreases also the hybridization rate. For a given MB the rates of hybridization decrease between one and two orders of magnitude when the stem length is increased from two to four nucleotides [27]. Shorter stems produce less stable hairpin structures, and in consequence reduce the signal background ratio and the selectivity of the MB-target interaction. For an optimal stem design, it is necessary to determine the balance between selectivity and hybridization rate for a given assay [26, 27, 68]. A typical MB has a stem of 5–7 nucleotides and a loop of 15–25 nucleotides in length (Figures 2(a) and 2(b)) and hybridize its target using only the loop region (Figure 2(b)). In order to accelerate the hybridization rate for MBs with more stable stems, one can increase the loop length, therefore augmenting the region of the MB that hybridize to the target. Another possibility, which prevents increasing the length of the loop, but ensures faster hybridization rates, is to include the stem sequence as part of the complementary region that will hybridize to the target. The MB could hybridize the target using one arm of the stem totally (Figure 2(c) for the 3′ arm or Figure 2(d) for the 5′ arm) or partially, or using partially both arms of the stem (Figure 2(e)). MBs using totally one arm of the stem to hybridize the target are known as shared stem (Figures 2(c) or 2(d)) [57]. For a given target sequence one shared-stem MB offers a high hybridization rate in comparison with a MB with equal number of complementary nucleotides, using only the loop as hybridization region, this is due to the complementary-nucleotides/MB-length ratio

$$\text{Hybridization rate} \propto \frac{\text{complementary nucleotides}}{\text{MB-length}}. \quad (1)$$

For living cells assays the MB have to discriminate among perfect matched and single mismatched targets at 37°C, it means that the melting temperature of the MB-single-mismatched target pair have to be less than 37°C, while the melting temperature of the MB-perfect-matched target pair must be above 37°C [69]. This requirement could be covered during the design of the stem stability, loop length, and/or regions that will participate in the hybridization. If the MB is designed with too high MB-target melting temperatures, it will be impossible to differentiate between the perfect complementary and the single mismatch targets. On the other

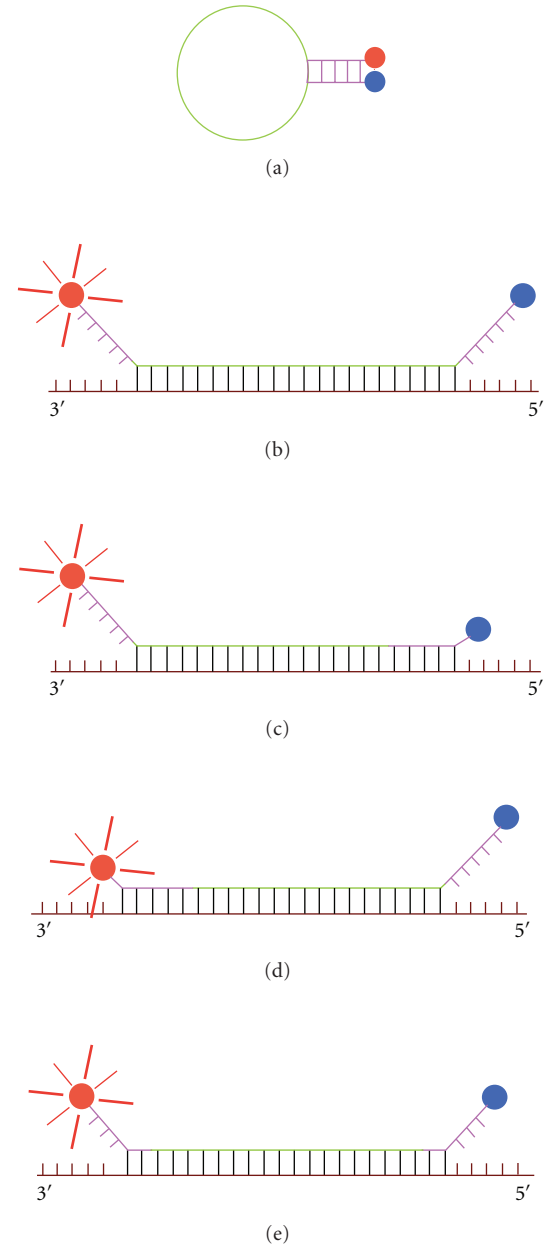


Figure 2: Different positions for MB-target hybridization. The loop region is illustrated in green and the stem in purple for all panels, target is brown and the bonds MB target are black ((b), (c), (d) y (e)). (a) MB with a stem of five nucleotides in the stem-loop hairpin conformation. (b) MB with a stem of five nucleotides and twenty two nucleotides in the loop, target hybridization occurs only at loop region. (c) MB with a stem of five nucleotides and a stem of seventeen nucleotides, using the loop and all the 3′ arm to hybridize the target (3′ shared stem MB). (d) MB with a stem of five nucleotides and a loop of seventeen nucleotides, using the loop and completely the 5′ arm to hybridize its target (5′ shared stem MB). (e) MB with a stem of five nucleotides and loop of eighteen nucleotides using partially both arms to hybridize its target (two nucleotides of every arm). Notice that a target that makes possible the design of a MB that hybridizes it using completely both arms could have a strong secondary structure that makes impossible the MB-target hybridization.

hand, if the MB-target melting temperatures were reduced too much, just a little fraction of the perfectly complementary target would be bound to the MB at physiological temperatures.

4. Chemistry Approaches for High Performance Molecular Beacons

When a MB will be used for RNA imaging in living cells, it must have minimal requirements in terms of sensitivity and selectivity, to prevent false-positive signals derived from nonspecific interactions, like protein binding and/or nuclease digestion. One could imagine that if the MB were digested by a nuclease, then the separation of the quencher and the reporter would necessarily result in fluorescence signal, which would not reflect MB-target interactions. Similarly, if protein binding diminishes quencher efficiency, fluorescence increments will not reflect the expected MB-target association.

4.1. MB with Chemically Enhanced Sensitivity. Some RNA species have a very low number of copies in living cells (e.g., microRNAs), which represent a big challenge for detection using MB. In such case, it is necessary to build MB with enhanced sensitivity and increased signal to background ratio. Increasing the efficiency of the quenching or increasing the fluorescence intensity emitted by the reporter after target hybridization could accomplish this. Most of the recent strategies are focused on increasing the fluorescence intensity of the reporter.

A typical quenching efficiency by contact or static quenching is in the range of 85–97% [70], depending on the factors previously discussed. Increasing the quenching efficiency would necessarily result in increased signal to background ratios. The increment in the dynamic range increases the difference of the fluorescence emitted by a population of MB-perfect-matched target hybrids in comparison with an equivalent population of MB-single-mismatched targets.

Recently developed MB synthesis strategies to increase the quenching efficiency consist in adding several quencher molecules in tandem (the so-called superquenching). Recent MBs have been synthesized using a FAM reporter and two or three Dabcyl molecules as quenchers (Figure 3) at the 5′ end [19]. This strategy results in an increment in quenching efficiency from 92.9% for one Dabcyl to 98.75% for two Dabcyl and 99.7% for three Dabcyl quenchers. Also the signal enhancement increases from 14 (one Dabcyl) to 80 (two Dabcyl) and 320 units (three Dabcyl molecules), increasing significantly the chances of detecting very low quantities of target [19]. Another interesting observation resulting from the use of three Dabcyl in tandem is that the increase in the hydrophobic interactions that occurs between the reporter and quenchers cause that the melting temperature of the MB stem increases in 4°C in comparison with stems having a single Dabcyl quencher. In theory, the higher melting temperature could enhance the capability of the MB to discriminate between perfect matched and single mismatched targets.

Since fluorescence quenching is based on static and dynamic quenching mechanisms and the efficiency of both depends on molecular distances and spatial orientation (dipole moment), having multiple quenchers in tandem may facilitate that one or more of them attain the correct spatial position in respect to the reporter, thus explaining the high efficiency of the super-quenching phenomenon.

Some polymeric fluorescent dyes could be employed to increase the fluorescence intensity and also the sensitivity of MB. Poly(phenylene ethynylene) (PPE) [71] is a water soluble polyelectrolyte with a high quantum yield [72]. Using a novel polymerization reaction it is possible to couple PPE directly to oligonucleotides. PPE have the highest fluorescent intensity compared to common fluorescent dyes used in MB synthesis. PPE has a fluorescent intensity around 20-fold higher than Cy3 and more than 6-fold in comparison with Alexa Fluor 488 (the fluorescent dye with highest fluorescence intensity currently used for MB building). In addition the fluorescence intensity of a single PPE chain is about 75% of the brightness intensity obtained with a quantum dot. The Dabcyl quencher shows good quenching performance with PPE. Nevertheless, super-quenching with Dabcyl could increase the performance of a PPE-based MB and enhance the signal to a level that may facilitate the identification of low abundance RNA molecules in living cells.

4.2. Nuclease Resistant MB. Cells use nucleases for nucleic acid catabolism and reuse the nucleotides for nucleic acid synthesis. Nucleases are used for cellular defense against foreign nucleic acids and for degradation of damaged DNA. Unfortunately for the researcher, nucleases also depredate MB [73, 74], which constitutes a problem for RNA imaging in living cells using these probes. Recent experiments show that living cells exhibit increments in fluorescence intensity after 45 minutes of MB delivery, even in the absence of a target [75]. Data indicate that this fluorescence increment is due to MB nuclease-mediated degradation. To produce nuclease resistant MB with improved stability at the cytoplasm, modified nucleotides have been incorporated in MB synthesis. These modified nucleotides include 2′-O-methylated [76–78], phosphorothioate derivatives [79, 80], peptide nucleic acids (PNA) [81, 82], and locked nucleic acids (LNA) [83]. 2′-O-methylated MBs offer good nuclease resistance and also resist RNase activity [76]; the main disadvantage for its use is the high background due to nonspecific interactions with proteins [84–86], another common problem is probe accumulation at the nucleus [87] and mitochondria [88]. In order to avoid the nuclear accumulation a quantum dot (QD) has been recently linked to the 2′-O-methylated MB [89], the linker is a biotin-dT group at the 3′ stem. The slow accumulation of the 2′-O-methylated MB at the mitochondria is reported only when using cyanine labeled MBs. PNA-based MBs have good affinity for both DNA and RNA targets in conjunction with high resistance to nuclease degradation. However, their reduced solubility and occasional aggregation limits its use for in vivo localization studies. LNA have a bicyclic furanose unit locked in an RNA-mimicking sugar conformation. LNA-based MBs show higher affinity for its target than DNA-based MBs; also the hybridization rate is slower in comparison. The use of chimeric DNA-LNA MBs results in nuclease resistant probes,

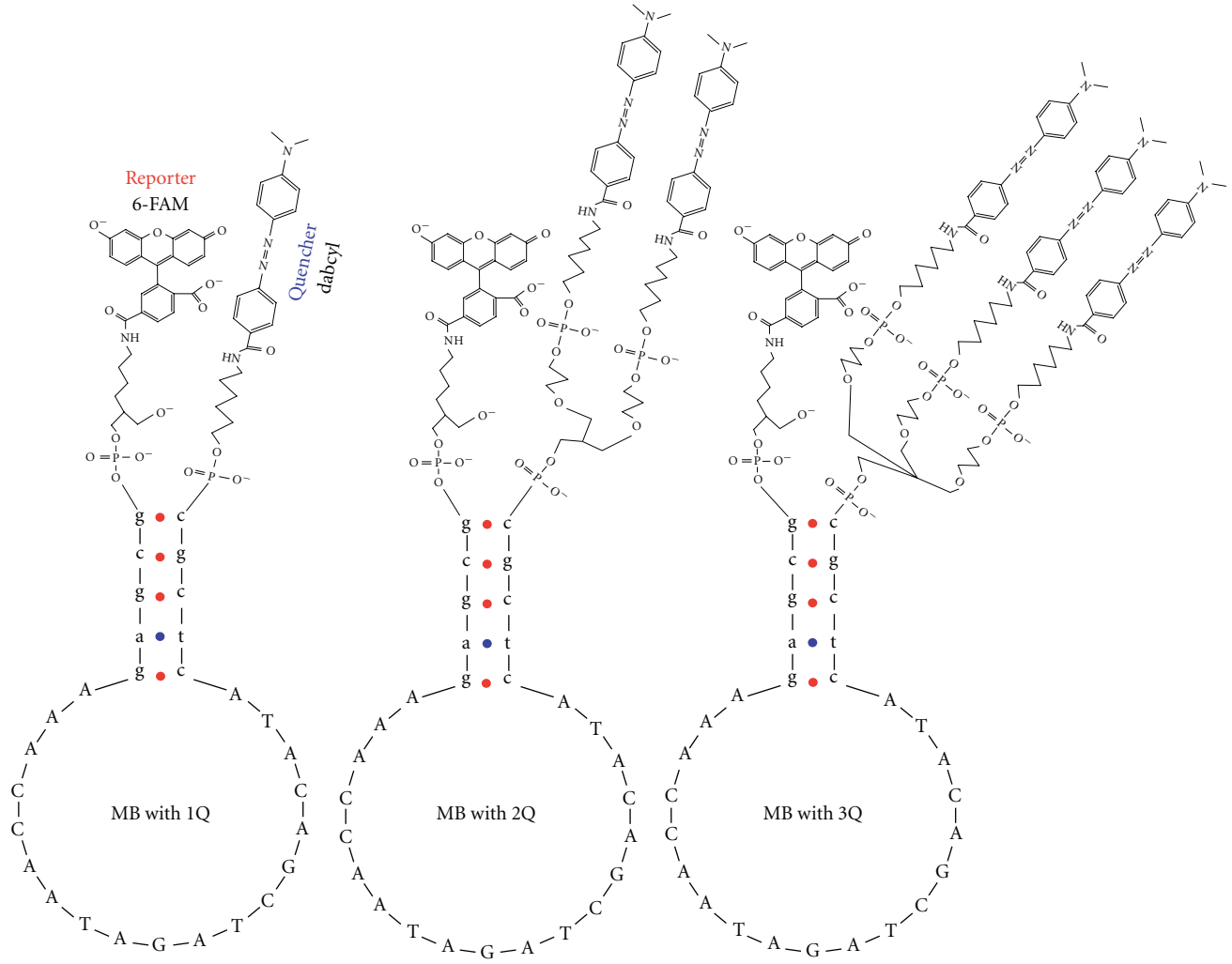


Figure 3: Structures of the superquenched MB. MB with one (1Q), two (2Q), and three (3Q) Dabcyl quenchers. The structures are represented in their oxidized state. The linker structures, the reporter linked to the 3′ end and the quencher or quenchers attached at the 5′ terminus correspond to [19].

with very high selectivity. It is important to take in consideration the LNA/DNA ratio of the chimeric probes, and if LNA bases will be part of the stem or not. By lowering the DNA/LNA ratio the hybridization rate is increased but also the nuclease degradation [83]. If the LNA bases are located at the stem, MB stability increases significantly, improving the selectivity. A shared-stem design with a four basepair stem and alternating DNA/LNA bases (ratio 1 : 1) has been recently suggested for living cell applications [90]. With this design a reasonable hybridization rate is obtained in addition to lower unspecific protein interactions and high nuclease resistance.

A novel strategy to avoid the possibility of the stem interactions with other nucleic acids or MB-MB between sticky ends consists in the use of L-DNA (specular isomer for the D-DNA) in the stem region [91]. D-DNA cannot interact with L-DNA and form left handed double helix. This design increases stem stability and MB selectivity.

5. Methods for MB Delivery into Cells

The efficient delivery of the MB into living cells is not an easy task. The method used must be efficient in terms that sufficient probe must be introduced in order to detect even low abundance RNA. Many delivery methods have been explored for MB introduction including microinjection, reversible cell membrane permeabilization, electroporation, MB linked to cell-penetrating peptides, and bioballistics. Needless to say that there is not a single method that can be used in all experimental conditions, therefore method selection is also based on the sample utilized.

Many transfection agents produce punctuate fluorescent patterns due to the passage of the MB by the endocytic pathway. Experimental evidence shows that linear fluorescently labeled probes enter in the endosomal/lysosomal pathway with the concomitant nuclease-mediated degradation. Reduced amounts of the probe may escape from the endocytic

pathway (0.01–10%) after several hours. For these reasons a delivery method in which the MB could enter the cell cytoplasm evading the endosomal pathway is highly desirable (e.g., microinjection, permeabilization of cell membranes by toxins, MB linked to cell penetration peptides, electroporation, or bioballistics).

5.1. Microinjection. The microinjection [42, 92] is in fact the most invasive method for MB delivery into living cells, which may result in cell damage and/or cell death. In the case where cells do not die, the injection could interfere with normal cellular functions and also produce altered or not reproducible results. An other disadvantage is the low number of cells that could be injected, given the fact that one must inject cell by cell. The major advantage is that low MB volumes are used reducing the cost of MB delivery. When using this method it is important to include a nonfluorescent dye (ink) in the solution, to ensure that the small volume enters the cell. Ink selection is also relevant to reduce side effects. Perhaps the main constrain for using this method is that a specialized injection system must be used in order to deliver a few nanoliters (nL) into the cell. Typically mammalian cells cannot be injected with volumes larger than 5 nL without compromising cell viability.

5.2. Toxin-Mediated Cell Membrane Permeabilization. A nonendocytic delivery method is the reversible membrane permeabilization mediated by toxins. Streptolysin O (SLO) is a bacterial toxin produced by streptococci hemolytic strain (*Streptococcus* pyogenes) having the ability to form pores on the cell membrane [93–95]. SLO binds as a monomer to cholesterol located on the cell plasma membrane, later SLO oligomerizes into a ring structure, producing a transient pore with a diameter in the range of 25–30 nm. The pore allows the entrance of macromolecules like the MB. The SLO cell membrane permeabilization method must be conducted in serum-free conditions; after membrane permeabilization the addition of the culture media containing serum inactivates the toxin [93, 96]. Since cholesterol composition is not a constant among cell types, the permeabilization process have to be optimized for every cell type, testing the MB entrance efficiency under different conditions of temperature, SLO concentration, cell number, and incubation time. We have found the use of combined MB-target solutions extremely useful for such testing to visualize the increment of cytosolic fluorescence as the MB-target enter the cell via the SLO pore.

5.3. MB Linked to Cell-Penetrating Peptides. There are several naturally occurring peptides with the capacity to penetrate cellular membranes (CPP: cell-penetrating peptides) [97–99]. One of the most widely used is, for instance, the transactivating transcriptional activator (Tat) from human immunodeficiency virus (HIV-1). Another peptide is a fragment of the Drosophila antennapedia homeodomain (RQIKIWFQNRRMKWKK) and the peptide VP22 from the tegument of the herpes simplex virus type 1 (HSV-1). In general, CPPs are short peptides rich in basic amino acids presenting an amphipathic arrange. CPPs penetrate across the cellular membranes without toxic effects and with very high efficiency (near to 100% of cells exposed are affected). The mechanism underlying CPP internalization prevents membrane destabilization, loss of cellular integrity and appears to be independent from the endocytic pathway (therefore reducing MB degradation by nuclease attack). Currently the most widely used CPP is the HIV-1 Tat peptide, and also the best studied due to its small size and high delivery efficiency.

CPPs deliver a wide variety of cargo molecules in culture cells or tissues [100, 101]. A major advantage of CPPs is that they do not interfere with specificity or sensitivity of MBs.

Recently CCP and NLS peptides (peptides with nuclear localization sequences) were conjugated to MB by different linkers (being the most common the carbon saturated chains (Figures 5(b), 5(b′) and 5(c)). The maleimide-thiol system is a very useful tool to conjugate MB with peptides. Maleimide reacts with moieties having thiol groups (Figure 4(a)), like cysteine or free thiol groups. The first option is to have a sulfhydryl in the terminus of the MB linker to react with maleimide at the terminus of the peptide (Figure 4(b)). The second option is the opposite, in that the MB linker has the maleimide and the peptide has the thiol group (Figure 4(b′)). A third option consists in having two sulfhydryl groups, one at the terminus of the MB linker and another at the terminus of the peptide (Figure 4(c)) forming a disulphide bond. This last option is the most versatile, because it facilitates the separation of the MB and the peptide using a reducing environment. Using CCPs as tools made possible imaging mRNAs like GAPDH recently [102]. Another interesting strategy involves the linking of the MB to peptides with nuclear localization signals (NLSs), in conjunction with SLO facilitates reaching nuclear targets like U1, U2, or U3 [103] (Figure 4(d)).

5.4. Electroporation. Electroporation methods have been used for delivering MBs into the cytoplasm of living cells, because electroporation avoid the endosomal pathway and in consequence reduces MB degradation. In the past, electroporation was associated with low cellular viability, but recent advances in this technology have led to a reduction in the events associated with cell damage or stress, including heat generation, metal ion dissolution, pH variations, or free radicals generation. Recently a novel electroporation method called microporation shows that a short time after the electroporation process occurs, a uniform distribution of the probe is observed at the cytoplasm [104]. Using this novel method high delivery efficiencies superior to 90% and cell viability near to 86% are observed [104]. With microporation (or any electroporation method) cells must be in suspension during the delivery procedure, therefore a time is required for cell to adhere before imaging studies can be conducted. This waiting period prevents the observation of early events of RNA mobilization or transport, a limitation to be taken into account during experiment planning.

5.5. Bioballistics. In the bioballistics technique, heavy metal particles are coated with molecules (e.g., MB) and propelled into the cells via a gas gun. The first application of this

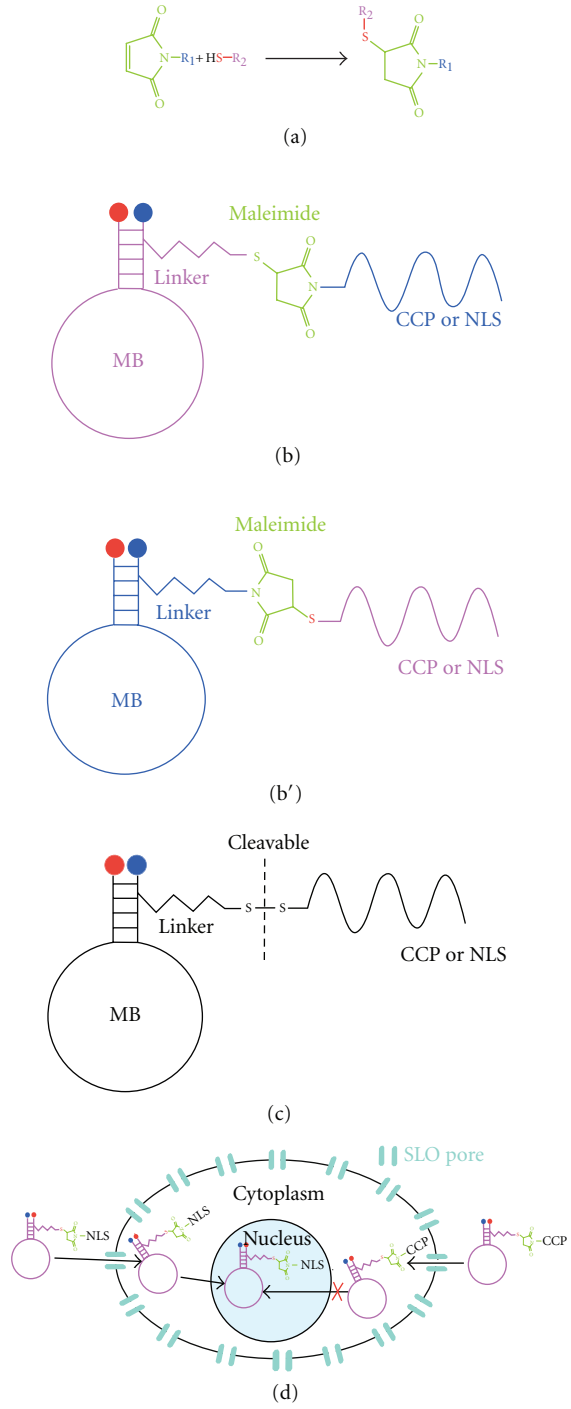


Figure 4: MB linked to penetrating peptides. (a) Reaction of conjugation among the maleimide molecule with a carrier group (R_1) and the thiol group of another molecule. This reaction may possibly link MB-peptide and occurs at a pH between 6.5–7.5. (b) One possibility using the maleimide-thiol system is that the peptide (NLS or CCP) it is linked to the maleimide across the secondary amine and reacts with a sulfhydryl group at the terminus of the MB hydrocarbon linker. (b′) The other possibility for the maleimide-thiol system is that the maleimide is linked at the terminus of the MB hydrocarbon linker and reacts with a thiol group at the peptide (NLS or CCP). (c) The link between the hydrocarbon linker of the MB and the peptide also could be across a disulphide bridge, in this case exists the possibility to cleave the bond in a reducing environment. (d) Using SLO to make permeable the cytoplasmic membrane is possible to introduce a MB linked to an NLS peptide to the cytoplasm with the objective that it be transported into the nucleus by the cellular machinery. For a CCP-linked MB the entrance to the nucleus is impossible.

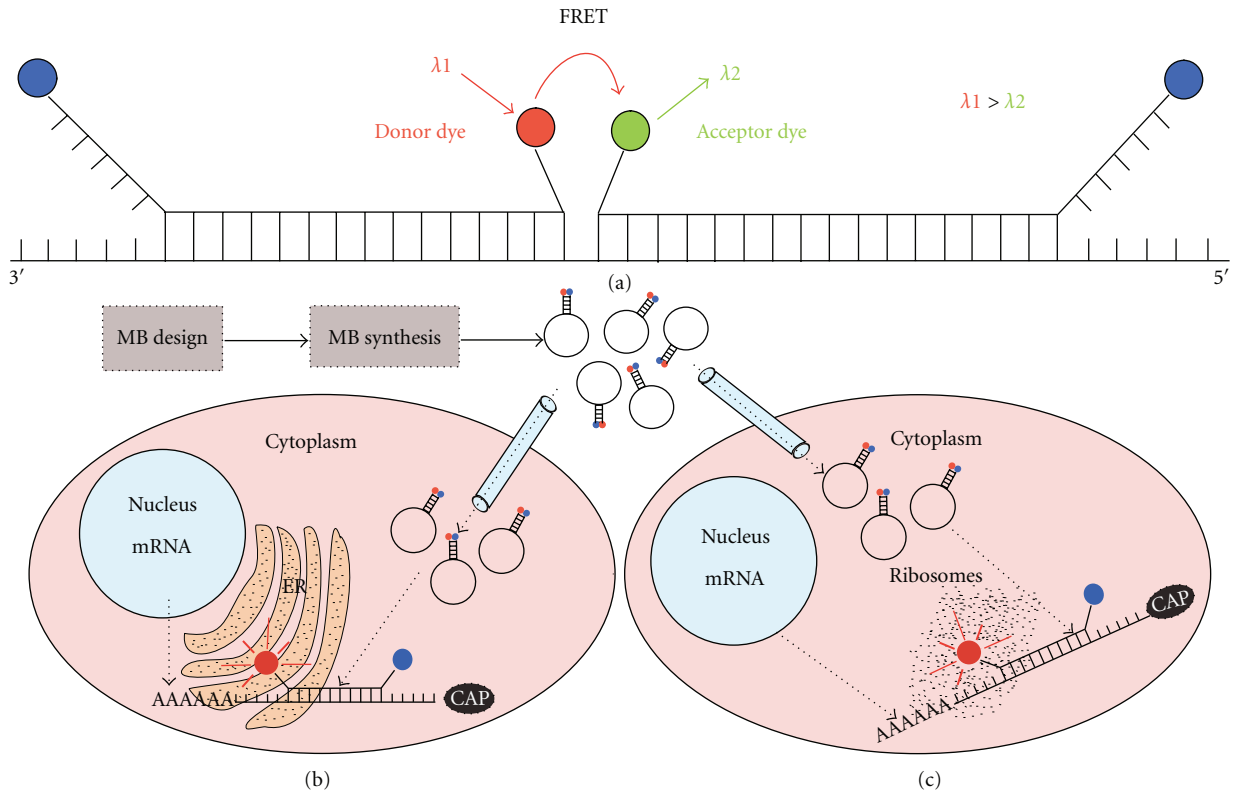


Figure 5: Approaches for RNA visualization in living cells. (a) In order to reduce the possibility of false-positive signals and also increase the specificity, one can design two shared stem MB having a pair reporters compatible with FRET, one to serve like a donor dye (red circle) and the other like a acceptor dye (green circle). With this strategy the sequence to be recognized increases its length and thus the specificity of the assay (observe that one MB needs to have the reporter at the 3′ end and the quencher at the 5′ end, in a not conventional attachment). After the MB design and synthesis the probe could be delivered using microinjection; in conjunction with other MB or fluorescent dyes could be used to measure RNA distribution at subcellular level. For example it is possible to use a specific fluorescent dye for the endoplasmic reticulum (ER) and also a MB designed to hybridize a given sequence of the ribosome. If a particular mRNA (based on the fluorescence of the MB) colocalizes with the ER dye and ribosome dye it means that this mRNA is translated in ribosomes at ER (b) or if the mRNA signal colocalizes only with the ribosome dye means that the mRNA translation occurs at free ribosomes (c); also it is possible that the mRNA translation takes place at both free and ER associated ribosomes. These are only some examples of the power of using MB in combination with other compartment or structure-specific dyes.

technique was in 1987 on plant cells [105]. In 1990 the technique was used for the first time for nucleic acids delivery into mammalian cells [106]. The biobullets usually are made with biocompatible metals, for example, gold or tungsten, having diameters from 1 to 5 micrometers [107]. Bullets must be coated with nucleic acids of interest prior to loading into the gun. The particles are propelled by the discharge of a gas into the gun chamber at very high pressure, usually helium [108] or water vapor [106]. The particles enter into the cytoplasm independently of the endosomal pathway, thus avoiding degradation [109]. A disadvantage of this method is that the cell membrane could present alterations that may induce cell stress. Bio-ballistic techniques are less invasive than microinjection and membrane alterations associated with the shot are usually short lasting. Perhaps the major disadvantage for bioballistics is that there is no control over the amount of MB delivery into the cell. Also the amount of MB entering each cell is variable. This method is most suited for cell population studies, but may hamper the interpretation of single cell imaging studies.

6. Real-Time Monitoring of RNA in Living Cells Using MB

6.1. mRNA Imaging in Living Cells Using MB. MBs have been only recently used for monitoring mRNA in living cells. MBs are particularly suited for mRNA studies because these RNA molecules are usually in high abundance, or at least higher than noncoding RNAs, for instance. For a successful MB design one needs to take into consideration all the factors involved in mRNA recognition by MB. Perhaps the most complicated issue is the selection of an appropriate target region in the mRNA sequence. One must first analyze the predicted secondary and tertiary structures in the mRNA target in an attempt to identify those regions showing minimum structure [68, 110]. In general, heavy secondary structure hampers MB binding to the target at room temperature, or results in very slow association kinetics. One should always consider that the determination of the target secondary structure is based on *in silico* predictions, which may or may not reflect the real structure of that region selected *in vivo*.

Another relevant consideration is to ensure that the region selected is not present in other nucleic acid in the cell. This can be partially accomplished using alignment tools like Blast [111], using the MB sequence as input. To ensure that the MB hybridizes with the expected target, a clever strategy has been recently implemented. In this strategy, the use of two shared stem MBs (one 3′ shared stem and the other 5′ shared stem) in adjacent positions is required (Figure 5(a)), both MBs are designed as FRET pair reporters to eliminate the possibility that false positives detection occurs [112, 113]. Shared stem design reduces the degree of freedom in the movements of both components of the FRET pair, enhancing FRET efficiency. Using two MBs reduces significantly the possibility of false positives, since having adjacent sequences in a given RNA that is recognized by two independent MBs is equivalent to recognizing a larger sequence (the sum of the MB-target complementary nucleotides in every MB).

Another promising application for MB research involves its use in conjunction with fluorescent dyes for specific organelles or subcellular structures. For example an assay designed to identify if a given mRNA is translated at ribosomes associated with endoplasmic reticulum (ER) or in free ribosomes could be developed, using a selective fluorescent dye for the ER and a MB designed to hybridize with the ribosomes and a second MB designed to detect the mRNA. With these elements it is possible to observe the colocalization of the mRNA within the ER and ribosomes simultaneously (Figure 5(b)) or visualize the colocalization of the mRNA with ribosomes only (no ER labeling) if the translation occurs at free ribosomes (Figure 5(c)).

6.2. Monitoring Gene Expression in Living Cells. The first assays where MB were involved focused only in detecting synthetic targets in *in vitro* studies, but in 2003 Bratu et al. demonstrated that a MB constitutes a powerful tool for RNA imaging and could be used for visualizing the distribution and transport of mRNA in living cells [48]. To reduce the background noise and ensure that the MB hybridizes its target, two MBs that recognize adjacent positions at the target were designed. The binding of both MBs resulted in FRET signal (energy transfer from one MB to the adjacent MB), indicating that both MBs were bound to adjacent regions in the target [112, 113]. The robustness of this method allows not only the identification of the mRNA target localization in the cell, but also opens the possibility to follow the transport of the mRNA through the cell in real time. Diffusion coefficients of the mRNA can be accurately measured combining fluorescence recovery after photobleaching (FRAP). In addition Santangelo et al. were beyond and analyzed the distribution and transport of the mRNA with respect in organelles from human fibroblasts. For this study MBs were used in combination with fluorescent stains of mitochondria, using spectral resolution Santangelo et al. demonstrated that both mRNA for glyceraldehydes 3-phospahate dehydrogenase and K-ras were specially enriched within the mitochondria [112]. Recently Wang et al. described a method to determine the expression levels of a specific RNA using MBs [114]. They use a system with two MBs and a reference probe as an internal standard for ratiometric analysis. The reference probe was used to avoid the variations due to different MB concentrations in the cell. One MB was used to detect the desired mRNA whereas the other MB was used to hybridize a house keeping mRNA, serving as reference. Three different cell types were used to compare expression levels of the mRNA of interest with excellent results.

Even though several methods for the study of unique mRNA molecules have been implemented, monitoring multiple RNA molecules in the same cell remains a technical challenge. The ratiometric analysis offers a good approach for the quantification of gene expression [115], enabling a more accurate estimation of mRNA levels. Even though in this particular study two different MBs were used, one can envision an extension of this method to study several RNA molecules simultaneously using spectrally different fluorescent reporters for each MB.

6.3. Monitoring Viral Genomic RNA Replication in Living Cells. RNA viruses have a genome composed by RNA with structural differences, which allow their classification into groups. Three groups of RNA viruses have been described, group III with a segmented double-stranded RNA genome, group IV with positive-sense single-stranded RNA genome, and group V with a negative-sense single-stranded genome. Methods for live-cell imaging of viral genomic RNA replication using MBs are currently available. Recently a method was implemented to monitor virus replication in living cells with the B6 coxsackievirus [116]. Virus replication was monitored using 2′-O-methylated RNA-MB. MB was introduced in the cells using a thiol-maleimide bridge linked to the Tat peptide. The MB was directed to a sequence of 18 basepairs located in the 5′ noncoding region of the viral genome. Coxsackieviruses belong to a Picornaviridae family having a positive-sense single-stranded genome. In a different report, researchers monitored genomic RNA replication of the bovine respiratory syncytial virus (bRSV) [46] using a MB directed to tandem-repeated consensus-sequences at intergenic regions, in order to amplify the signal. Detection of viral genomic RNA in living cells is a promising application because it facilitates the detection of viruses without the need for RNA purification, cDNA synthesis steps and/or amplification procedures, and also open the possibility of obtaining spatial-temporal information contributing to the understanding of viral replication, delivery, or propagation mechanisms. These contributions could facilitate the development of drugs for viral infection treatment or rapid diagnostics techniques. Undoubtedly, MBs will impulse viral research by facilitating the study of genome replication and transport in living cells, something difficult to envision before the MB era.

7. Conclusion

We have discussed in this paper important advances in the design, synthesis, and use of MBs since its introduction by Tyagi and Kramer in 1996. We have highlighted relevant issues involved in sequence selection, loop, and stem considerations needed for a successful MB design. We

have mentioned also new developments in MB synthesis using modified nucleic acids highlighting its advantages and limitations. Finally, we have gathered recent studies where MBs have been used in rather clever experiments directed to the study of RNA localization and dynamics in living cells.

The development of novel fluorescent indicators with greater dynamic range in combination with robust quenchers, in the next generation of MBs, will provide exceptional tools for RNA studies in living cells and tissues. The combination of such tools with powerful imaging techniques, such as FRAP, FLIM, TIRFM, and multiphoton confocal microscopy promises the possibility to conduct such studies in living organisms.

Finally, a poorly explored area in MB research is the use of these tools in diagnostics. One can easily envision methodologies capable of identifying replicating viruses in human samples, using MB in combination with fluorescent assays without the need of time consuming more elaborated techniques such as qPCR.

Acknowledgments

This work was partially supported by a grant from the Instituto de Ciencia y Tecnología del Distrito Federal (ICYTDF) and Dirección General de Asuntos del Personal Académico (DGAPA) to L. Vaca. R. Monroy-Contreras is a fellow from the Consejo Nacional de Ciencia y Tecnologia (CONACyT). The technical assistance from Alicia Sampieri is greatly appreciated.

References

[1] V. Glišin, R. Crkvenjakov, and C. Byus, "Ribonucleic acid isolated by cesium chloride centrifugation," *Biochemistry*, vol. 13, no. 12, pp. 2633–2637, 1974.

[2] C. Auffray and F. Rougeon, "Purification of mouse immunoglobulin heavy-chain messenger RNAs from total myeloma tumor RNA," *European Journal of Biochemistry*, vol. 107, no. 2, pp. 303–314, 1980.

[3] P. Chomczynski and N. Sacchi, "Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction," *Analytical Biochemistry*, vol. 162, no. 1, pp. 156–159, 1987.

[4] R. K. Saiki, S. Scharf, and F. Faloona, "Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia," *Science*, vol. 230, no. 4732, pp. 1350–1354, 1985.

[5] J. C. Alwine, D. J. Kemp, B. A. Parker et al., "Detection of specific RNAs or specific fragments of DNA by fractionation in gels and transfer to diazobenzyloxymethyl paper," *Methods in Enzymology*, vol. 68, pp. 220–242, 1979.

[6] V. E. Velculescu, L. Zhang, B. Vogelstein, and K. W. Kinzler, "Serial analysis of gene expression," *Science*, vol. 270, no. 5235, pp. 484–487, 1995.

[7] L. Diatchenko, Y. F. C. Lau, A. P. Campbell et al., "Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, no. 12, pp. 6025–6030, 1996.

[8] P. Liang and A. B. Pardee, "Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction," *Science*, vol. 257, no. 5072, pp. 967–971, 1992.

[9] N. Lisitsyn and M. Wigler, "Cloning the differences between two complex genomes," *Science*, vol. 259, no. 5097, pp. 946–951, 1993.

[10] M. D. Adams, M. Dubnick, A. R. Kerlavage et al., "Sequence identification of 2,375 human brain genes," *Nature*, vol. 355, no. 6361, pp. 632–634, 1992.

[11] M. Schena, D. Shalon, R. W. Davis, and P. O. Brown, "Quantitative monitoring of gene expression patterns with a complementary DNA microarray," *Science*, vol. 270, no. 5235, pp. 467–470, 1995.

[12] G. J. Bassell, C. M. Powers, K. L. Taneja, and R. H. Singer, "Single mRNAs visualized by ultrastructural in situ hybridization are principally localized at actin filament intersections in fibroblasts," *Journal of Cell Biology*, vol. 126, no. 4, pp. 863–876, 1994.

[13] M. C. Vaccaro, S. Gigliotti, F. Graziani et al., "A transient asymmetric distribution of XNOA 36 mRNA and the associated spectrin network bisects Xenopus laevis stage I oocytes along the future A/V axis," *European Journal of Cell Biology*, vol. 89, no. 7, pp. 525–536, 2010.

[14] Y. Shinmyo, T. Mito, T. Matsushita et al., "Caudal is required for gnathal and thoracic patterning and for posterior elongation in the intermediate-germband cricket Gryllus bimaculatus," *Mechanisms of Development*, vol. 122, no. 2, pp. 231–239, 2005.

[15] R. Krause, M. Hemberger, M. Messerschmid et al., "Molecular cloning and characterization of murine Mpgc60, a gene predominantly expressed in the intestinal tract," *Differentiation*, vol. 63, no. 5, pp. 285–294, 1998.

[16] M. Mione, Z. Lele, C. T. Kwong, M. L. Concha, and J. D. Clarke, "Expression of pcp4a in subpopulations of CNS neurons in zebrafish," *Journal of Comparative Neurology*, vol. 495, no. 6, pp. 769–787, 2006.

[17] S. Behrens, B. M. Fuchs, F. Mueller, and R. Amann, "Is the in situ accessibility of the 16S rRNA of Escherichia coli for Cy3-labeled oligonucleotide probes predicted by a three-dimensional structure model of the 30S ribosomal subunit?" *Applied and Environmental Microbiology*, vol. 69, no. 8, pp. 4935–4941, 2003.

[18] G. Bao, J. R. Won, and A. Tsourkas, "Fluorescent probes for live-cell RNA detection," *Annual Review of Biomedical Engineering*, vol. 11, pp. 25–47, 2009.

[19] C. J. Yang, H. Lin, and W. Tan, "Molecular assembly of superquenchers in signaling molecular interactions," *Journal of the American Chemical Society*, vol. 127, no. 37, pp. 12772–12773, 2005.

[20] S. Tyagi and F. R. Kramer, "Molecular beacons: probes that fluoresce upon hybridization," *Nature Biotechnology*, vol. 14, no. 3, pp. 303–308, 1996.

[21] S. D. Bélanger, M. Boissinot, C. Ménard, F. J. Picard, and M. G. Bergeron, "Rapid detection of Shiga toxin-producing bacteria in feces by multiplex PCR with molecular beacons on the Smart Cycler," *Journal of Clinical Microbiology*, vol. 40, no. 4, pp. 1436–1440, 2002.

[22] S. D. Bélanger, M. Boissinot, N. Clairoux, F. J. Picard, and M. G. Bergeron, "Rapid detection of Clostridium difficile in feces by real-time PCR," *Journal of Clinical Microbiology*, vol. 41, no. 2, pp. 730–734, 2003.

[23] A. Beloukas, D. Paraskevis, M. Psichogiou, and A. Hatzakis, "The role of HIV-1 DNA as an additional marker of HIV-1

infection," *Current HIV Research*, vol. 7, no. 3, pp. 255–265, 2009.

[24] M. G. Bergeron and D. Ke, "New DNA-based PCR approaches for rapid real-time detection and prevention of group B streptococcal infections in newborns and pregnant women," *Expert Reviews in Molecular Medicine*, vol. 3, no. 27, pp. 1–14, 2001.

[25] W. Chen, G. Martinez, and A. Mulchandani, "Molecular beacons: a real-time polymerase chain reaction assay for detecting Salmonella," *Analytical Biochemistry*, vol. 280, no. 1, pp. 166–172, 2000.

[26] G. Bonnet, S. Tyagi, A. Libchaber, and F. R. Kramer, "Thermodynamic basis of the enhanced specificity of structured DNA probes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 11, pp. 6171–6176, 1999.

[27] A. Tsourkas, M. A. Behlke, S. D. Rose, and G. Bao, "Hybridization kinetics and thermodynamics of molecular beacons," *Nucleic Acids Research*, vol. 31, no. 4, pp. 1319–1330, 2003.

[28] I. Tapp, L. Malmberg, E. Rennel, M. Wik, and A. C. Syvanen, "Homogeneous scoring of single-nucleotide polymorphisms: comparison of the 5'-nuclease TaqMan® assay and molecular beacon probes (BioTechniques (2000) 28 (732-738))," *BioTechniques*, vol. 28, no. 4, pp. 732–738, 2000.

[29] M. M. Mhlanga and L. Malmberg, "Using molecular beacons to detect single-nucleotide polymorphisms with real-time PCR," *Methods*, vol. 25, no. 4, pp. 463–471, 2001.

[30] S. A. Marras, F. R. Kramer, and S. Tyagi, "Genotyping SNPs with molecular beacons," *Methods in Molecular Biology*, vol. 212, pp. 111–128, 2003.

[31] L. B. Barreiro, R. Henriques, and M. M. Mhlanga, "High-throughput SNP genotyping: combining tag SNPs and molecular beacons," *Methods in Molecular Biology*, vol. 578, pp. 255–276, 2009.

[32] L. G. Kostrikis, S. Tyagi, M. M. Mhlanga, D. D. Ho, and F. R. Kramer, "Spectral genotyping of human alleles," *Science*, vol. 279, no. 5354, pp. 1228–1229, 1998.

[33] L. G. Kostrikis, S. Shin, and D. D. Ho, "Genotyping HIV-1 and HCV strains by a combinatorial DNA melting assay (COMA)," *Molecular Medicine*, vol. 4, no. 7, pp. 443–453, 1998.

[34] G. Orrù, G. Faa, S. Pillai et al., "Rapid PCR real-time genotyping of M-Malton α1-antitrypsin deficiency alleles by molecular beacons," *Diagnostic Molecular Pathology*, vol. 14, no. 4, pp. 237–242, 2005.

[35] G. Orru, F. Coghe, G. Faa et al., "Rapid multiplex real-time PCR by molecular beacons for different BRAF allele detection in papillary thyroid carcinoma," *Diagnostic Molecular Pathology*, vol. 19, no. 1, pp. 1–8, 2010.

[36] K. Szuhai, E. Sandhaus, S. M. Kolkman-Uljee et al., "A novel strategy for human papillomavirus detection and genotyping with sybrgreen and molecular beacon polymerase chain reaction," *American Journal of Pathology*, vol. 159, no. 5, pp. 1651–1660, 2001.

[37] B. A. J. Giesendorf, J. A. M. Vet, S. Tyagi, E. J. M. G. Mensink, F. J. M. Trijbels, and H. J. Blom, "Molecular beacons: a new approach for semiautomated mutation analysis," *Clinical Chemistry*, vol. 44, no. 3, pp. 482–486, 1998.

[38] K. Frei, K. Szuhai, T. Lucas et al., "Connexin 26 mutations in cases of sensorineural deafness in eastern Austria," *European Journal of Human Genetics*, vol. 10, no. 7, pp. 427–432, 2002.

[39] D. R. Hodgson, C. A. Foy, M. Partridge, S. Pateromichelakis, and N. J. Gibson, "Development of a facile fluorescent assay for the detection of 80 mutations within the p53 gene," *Molecular Medicine*, vol. 8, no. 5, pp. 227–237, 2002.

[40] J. B. Biggins, J. R. Prudent, D. J. Marshall, M. Ruppen, and J. S. Thorson, "A continuous assay for DNA cleavage: the application of 'break lights' to enediynes, iron-dependent and nucleases," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 25, pp. 13537–13542, 2000.

[41] J. J. Li, R. Geyer, and W. Tan, "Using molecular beacons as a sensitive fluorescence assay for enzymatic cleavage of single-stranded DNA," *Nucleic acids research*, vol. 28, no. 11, article e52, 2000.

[42] C. D. Medley, T. J. Drake, J. M. Tomasini, R. J. Rogers, and W. Tan, "Simultaneous monitoring of the expression of multiple genes inside of single breast carcinoma cells," *Analytical Chemistry*, vol. 77, no. 15, pp. 4713–4718, 2005.

[43] X. H. Peng, Z. H. Cao, J. T. Xia et al., "Real-time detection of gene expression in cancer cells using molecular beacon imaging: new strategies for cancer research," *Cancer Research*, vol. 65, no. 5, pp. 1909–1917, 2005.

[44] L. Yang, Z. Cao, Y. Lin, W. C. Wood, and C. A. Staley, "Molecular beacon imaging of tumor marker gene expression in pancreatic cancer cells," *Cancer Biology and Therapy*, vol. 4, no. 5, pp. 561–570, 2005.

[45] A. Wang, A. M. Salazar, M. V. Yates, A. Mulchandani, and W. Chen, "Visualization and detection of infectious coxsackievirus replication using a combined cell culture-molecular beacon assay," *Applied and Environmental Microbiology*, vol. 71, no. 12, pp. 8397–8401, 2005.

[46] P. Santangelo, N. Nitin, L. LaConte, A. Woolums, and G. Bao, "Live-cell characterization and analysis of a clinical isolate of bovine respiratory syncytial virus, using molecular beacons," *Journal of Virology*, vol. 80, no. 2, pp. 682–688, 2006.

[47] H. Y. Yeh, M. V. Yates, A. Mulchandani, and W. Chen, "Visualizing the dynamics of viral replication in living cells via Tat peptide delivery of nuclease-resistant molecular beacons," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 45, pp. 17522–17525, 2008.

[48] D. P. Bratu, B. J. Cha, M. M. Mhlanga, F. R. Kramer, and S. Tyagi, "Visualizing the distribution and transport of mRNAs in living cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 23, pp. 13308–13313, 2003.

[49] D. P. Bratu, "Molecular beacons light the way: imaging native mRNAs in living cells," *Discovery Medicine*, vol. 3, no. 19, pp. 44–47, 2003.

[50] S. Tyagi and O. Alsmadi, "Imaging native β-actin mRNA in motile fibroblasts," *Biophysical Journal*, vol. 87, no. 6, pp. 4153–4162, 2004.

[51] B. Simon, M. Sandhu, and K. L. Myhr, "Live FISH: imaging mRNA in living neurons," *Journal of Neuroscience Research*, vol. 88, no. 1, pp. 55–63, 2010.

[52] X. Fang, J. J. Li, J. Perlette, W. Tan, and K. Wang, "Molecular beacons: novel fluorescent probes," *Analytical Chemistry*, vol. 72, no. 23, pp. 747A–753A, 2000.

[53] W. Tan, X. Fang, J. Li, and X. Liu, "Molecular beacons: a novel DNA probe for nucleic acid and protein studies," *Chemistry*, vol. 6, no. 7, pp. 1107–1111, 2000.

[54] S. A. Marras, F. R. Kramer, and S. Tyagi, "Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes," *Nucleic Acids Research*, vol. 30, no. 21, article e122, 2002.

[55] K. M. Parkhurst and L. J. Parkhurst, "Donor- acceptor distance distributions in a double-labeled fluorescent oligonucleotide both as a single strand and in duplexes," *Biochemistry®*, vol. 34, no. 1, pp. 293–300, 1995.

[56] S. Bernacchi and Y. Mély, "Exciton interaction in molecular beacons: a sensitive sensor for short range modifications of the nucleic acid structure," *Nucleic Acids Research*, vol. 29, no. 13, article e62, 2001.

[57] A. Tsourkas, M. A. Behlke, and G. Bao, "Structure-function relationships of shared-stem and conventional molecular beacons," *Nucleic Acids Research*, vol. 30, no. 19, pp. 4208–4215, 2002.

[58] E. McGown, M. Su, and R. Dennis, "Measurement of molecular beacons in the SPECTRAmax® GEMINI spectrofluorometer," *Journal of Chemical Technology and Biotechnology*, vol. 75, no. 10, pp. 942–944, 2000.

[59] C. A. M. Seidel, A. Schulz, and M. H. M. Sauer, "Nucleobase-specific quenching of fluorescent dyes. 1. Nucleobase one-electron redox potentials and their Correlation with static and dynamic quenching efficiencies," *Journal of Physical Chemistry*, vol. 100, no. 13, pp. 5541–5553, 1996.

[60] T. Heinlein, J. P. Knemeyer, O. Piestert, J. Wolfrum, and M. Sauer, "Nucleobase-specific quenching of fluorescent dyes in DNA-hairpins," *The Journal of Physical Chemistry B*, vol. 107, no. 31, pp. 7957–7964, 2003.

[61] G. Luo, L. Zheng, X. Zhang, J. Zhang, P. Nilsson-Ehle, and N. Xu, "Genotyping of single nucleotide polymorphisms using base-quenched probe: a method does not invariably depend on the deoxyguanosine nucleotide," *Analytical Biochemistry*, vol. 386, no. 2, pp. 161–166, 2009.

[62] S. A. Oladepo and G. R. Loppnow, "Self-quenching smart probes as a platform for the detection of sequence-specific UV-induced DNA photodamage," *Analytical and Bioanalytical Chemistry*, vol. 397, no. 7, pp. 2949–2957, 2010.

[63] N. Venkatesan, Y. J. Seo, and B. H. Kim, "Quencher-free molecular beacons: a new strategy in fluorescence based nucleic acid analysis," *Chemical Society Reviews*, vol. 37, no. 4, pp. 648–663, 2008.

[64] P. M. Holland, R. D. Abramson, R. Watson, and D. H. Gelfand, "Detection of specific polymerase chain reaction product by utilizing the 5' → 3' exonuclease activity of Thermus aquaticus DNA polymerase," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, no. 16, pp. 7276–7280, 1991.

[65] S. Tyagi, S. A. E. Marras, and F. R. Kramer, "Wavelength-shifting molecular beacons," *Nature Biotechnology*, vol. 18, no. 11, pp. 1191–1196, 2000.

[66] S. Tyagi, D. P. Bratu, and F. R. Kramer, "Multicolor molecular beacons for allele discrimination," *Nature Biotechnology*, vol. 16, no. 1, pp. 49–53, 1998.

[67] L. Tan, Y. Li, T. J. Drake et al., "Molecular beacons for bioanalytical applications," *Analyst*, vol. 130, no. 7, pp. 1002–1005, 2005.

[68] M. Zuker, "Mfold web server for nucleic acid folding and hybridization prediction," *Nucleic Acids Research*, vol. 31, no. 13, pp. 3406–3415, 2003.

[69] J. Perlette and W. Tan, "Real-time monitoring of intracellular mRNA hybridization inside single living cells," *Analytical Chemistry*, vol. 73, no. 22, pp. 5544–5550, 2001.

[70] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer, New York, NY, USA, 3rd edition, 2006.

[71] C. J. Yang, M. Pinto, K. Schanze, and W. Tan, "Direct synthesis of an oligonucleotide-poly-(phenylene ethynylene) conjugate with a precise one-to-one molecular ratio," *Angewandte Chemie*, vol. 44, no. 17, pp. 2572–2576, 2005.

[72] H. Huang, K. Wang, W. Tan et al., "Design of a modular-based fluorescent conjugated polymer for selective sensing," *Angewandte Chemie*, vol. 43, no. 42, pp. 5635–5638, 2004.

[73] J. J. Li, R. Geyer, and W. Tan, "Using molecular beacons as a sensitive fluorescence assay for enzymatic cleavage of single-stranded DNA," *Nucleic Acids Research*, vol. 28, no. 11, article e52, 2000.

[74] C. J. Yang, J. J. Li, and W. Tan, "Using molecular beacons for sensitive fluorescence assays of the enzymatic cleavage of nucleic acids," *Methods in Molecular Biology*, vol. 335, pp. 71–81, 2006.

[75] C. J. Yang, C. D. Medley, and W. Tan, "Monitoring nucleic acids using molecular beacons," *Current Pharmaceutical Biotechnology*, vol. 6, no. 6, pp. 445–452, 2005.

[76] C. Molenaar, S. A. Marras, J. C. Slats et al., "Linear 2' O-Methyl RNA probes for the visualization of RNA in living cells," *Nucleic Acids Research*, vol. 29, no. 17, article e89, 2001.

[77] A. Tsourkas, M. A. Behlke, and G. Bao, "Hybridization of 2′-O-methyl and 2-deoxy molecular beacons to RNA and DNA targets," *Nucleic Acids Research*, vol. 30, no. 23, pp. 5168–5174, 2002.

[78] R. H. Kehlenbach, "In vitro analysis of nuclear mRNA export using molecular beacons for target detection," *Nucleic Acids Research*, vol. 31, no. 11, article e64, 2003.

[79] R. Shah and W. S. El-Deiry, "p53-dependent activation of a molecular beacon in tumor cells following exposure to doxorubicin chemotherapy," *Cancer Biology and Therapy*, vol. 3, no. 9, pp. 871–875, 2004.

[80] V. Vijayanathan, T. Thomas, L. H. Sigal, and T. J. Thomas, "Direct measurement of the association constant of HER2/neu antisense oligonucleotide to its target RNA sequence using a molecular beacon," *Antisense and Nucleic Acid Drug Development*, vol. 12, no. 4, pp. 225–233, 2002.

[81] K. Petersen, U. Vogel, E. Rockenbauer et al., "Short PNA molecular beacons for real-time PCR allelic discrimination of single nucleotide polymorphisms," *Molecular and Cellular Probes*, vol. 18, no. 2, pp. 117–122, 2004.

[82] H. Kuhn, V. V. Demidov, B. D. Gildea, M. J. Fiandaca, J. C. Coull, and M. D. Frank-Kamenetskii, "PNA beacons for duplex DNA," *Antisense and Nucleic Acid Drug Development*, vol. 11, no. 4, pp. 265–270, 2001.

[83] C. J. Yang, L. Wang, Y. Wu et al., "Synthesis and investigation of deoxyribonucleic acid/locked nucleic acid chimeric molecular beacons," *Nucleic Acids Research*, vol. 35, no. 12, pp. 4030–4041, 2007.

[84] X. Fang, J. J. Li, and W. Tan, "Using molecular beacons to probe molecular interactions between lactate dehydrogenase and single-stranded DNA," *Analytical Chemistry*, vol. 72, no. 14, pp. 3280–3285, 2000.

[85] R. Yang, J. Jin, Y. Chen et al., "Carbon nanotube-quenched fluorescent oligonucleotides: probes that fluoresce upon hybridization," *Journal of the American Chemical Society*, vol. 130, no. 26, pp. 8351–8358, 2008.

[86] N. Graf, M. Göritz, and R. Krämer, "A metal-ion-releasing probe for DNA detection by catalytic signal amplification," *Angewandte Chemie*, vol. 45, no. 24, pp. 4013–4015, 2006.

[87] A. K. Chen, M. A. Behlke, and A. Tsourkas, "Sub-cellular trafficking and functionality of 2′-O-methyl and 2′-O-methyl-phosphorothioate molecular beacons," *Nucleic Acids Research*, vol. 37, no. 22, article e149, 2009.

[88] W. J. Rhee and G. Bao, "Slow non-specific accumulation of 2'-deoxy and 2'-O-methyl oligonucleotide probes at mitochondria in live cells," *Nucleic Acids Research*, vol. 38, no. 9, article e109, 2010.

[89] A. K. Chen, M. A. Behlke, and A. Tsourkas, "Avoiding false-positive signals with nuclease-vulnerable molecular beacons in single living cells," *Nucleic Acids Research*, vol. 35, no. 16, article e105, 2007.

[90] Y. Wu, C. J. Yang, L. L. Moroz, and W. Tan, "Nucleic acid beacons for long-term real-time intracellular monitoring," *Analytical Chemistry*, vol. 80, no. 8, pp. 3025–3028, 2008.

[91] Y. Kim, C. J. Yang, and W. Tan, "Superior structure stability and selectivity of hairpin nucleic acid probes with an L-DNA stem," *Nucleic Acids Research*, vol. 35, no. 21, pp. 7279–7287, 2007.

[92] R. W. Dirks, C. Molenaar, and H. J. Tanke, "Visualizing RNA molecules inside the nucleus of living cells," *Methods*, vol. 29, no. 1, pp. 51–57, 2003.

[93] M. A. Barry and A. Eastman, "Identification of deoxyribonuclease II as an endonuclease involved in apoptosis," *Archives of Biochemistry and Biophysics*, vol. 300, no. 1, pp. 440–450, 1993.

[94] R. V. Giles, C. J. Ruddell, D. G. Spiller, J. A. Green, and D. M. Tidd, "Single base discrimination for ribonuclease H-dependent antisense effects within intact human leukaemia cells," *Nucleic Acids Research*, vol. 23, no. 6, pp. 954–961, 1995.

[95] R. V. Giles, D. G. Spiller, J. Grzybowski, R. E. Clark, P. Nicklin, and D. M. Tidd, "Selecting optimal oligonucleotide composition for maximal antisense effect following streptolysin O-mediated delivery into human leukaemia cells," *Nucleic Acids Research*, vol. 26, no. 7, pp. 1567–1575, 1998.

[96] I. Walev, S. C. Bhakdi, F. Hofmann et al., "Delivery of proteins into living cells by reversible membrane permeabilization with streptolysin-O," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 6, pp. 3185–3190, 2001.

[97] M. Becker-Hapak, S. S. McAllister, and S. F. Dowdy, "TAT-mediated protein transduction into mammalian cells," *Methods*, vol. 24, no. 3, pp. 247–256, 2001.

[98] E. L. Snyder and S. F. Dowdy, "Protein/peptide transduction domains: potential to deliver large DNA molecules into cells," *Current Opinion in Molecular Therapeutics*, vol. 3, no. 2, pp. 147–152, 2001.

[99] J. S. Wadia and S. F. Dowdy, "Protein transduction technology," *Current Opinion in Biotechnology*, vol. 13, no. 1, pp. 52–56, 2002.

[100] H. Brooks, B. Lebleu, and E. Vivès, "Tat peptide-mediated cellular delivery: back to basics," *Advanced Drug Delivery Reviews*, vol. 57, no. 4, pp. 559–577, 2005.

[101] J. S. Wadia and S. F. Dowdy, "Transmembrane delivery of protein and peptide drugs by TAT-mediated transduction in the treatment of cancer," *Advanced Drug Delivery Reviews*, vol. 57, no. 4, pp. 579–596, 2005.

[102] N. Nitin, P. J. Santangelo, G. Kim, S. Nie, and G. Bao, "Peptide-linked molecular beacons for efficient delivery and rapid mRNA detection in living cells," *Nucleic Acids Research*, vol. 32, no. 6, article e58, 2004.

[103] N. Nitin and G. Bao, "NLS peptide conjugated molecular beacons for visualizing nuclear RNA in living cells," *Bioconjugate Chemistry*, vol. 19, no. 11, pp. 2205–2211, 2008.

[104] A. K. Chen, M. A. Behlke, and A. Tsourkas, "Efficient cytosolic delivery of molecular beacon conjugates and flow cytometric analysis of target RNA," *Nucleic Acids Research*, vol. 36, no. 12, article e69, 2008.

[105] T. M. Klein, E. D. Wolf, R. Wu, and J. C. Sanford, "High-velocity microprojectiles for delivering nucleic acids into living cells," *Nature*, vol. 327, no. 6117, pp. 70–73, 1987.

[106] N. S. Yang, J. Burkholder, B. Roberts, B. Martinell, and D. McCabe, "In vivo and in vitro gene transfer to mammalian somatic cells by particle bombardment," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 87, no. 24, pp. 9568–9572, 1990.

[107] E. F. Fynan, R. G. Webster, D. H. Fuller, J. R. Haynes, J. C. Santoro, and H. L. Robinson, "DNA vaccines: protective immunizations by parenteral, mucosal, and gene- gun inoculations," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 90, no. 24, pp. 11478–11482, 1993.

[108] J. A. Wolff, P. Williams, G. Acsadi, S. Jiao, A. Jani, and W. Chong, "Conditions affecting direct gene transfer into rodent muscle in vivo," *BioTechniques*, vol. 11, no. 4, pp. 474–485, 1991.

[109] C. Condon, S. C. Watkins, C. M. Celluzzi, K. Thompson, and L. D. Falo, "DNA-based immunization by in vivo transfection of dendritic cells," *Nature Medicine*, vol. 2, no. 10, pp. 1122–1128, 1996.

[110] D. H. Mathews, J. Sabina, M. Zuker, and D. H. Turner, "Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure," *Journal of Molecular Biology*, vol. 288, no. 5, pp. 911–940, 1999.

[111] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool," *Journal of Molecular Biology*, vol. 215, no. 3, pp. 403–410, 1990.

[112] P. J. Santangelo, B. Nix, A. Tsourkas, and G. Bao, "Dual FRET molecular beacons for mRNA detection in living cells," *Nucleic Acids Research*, vol. 32, no. 6, article e57, 2004.

[113] A. Tsourkas, M. A. Behlke, Y. Xu, and G. Bao, "Spectroscopic features of dual fluorescence/luminescence resonance energy-transfer molecular beacons," *Analytical Chemistry*, vol. 75, no. 15, pp. 3697–3703, 2003.

[114] K. Wang, Z. Tang, C. J. Yang et al., "Molecular engineering of DNA: molecular beacons," *Angewandte Chemie*, vol. 48, no. 5, pp. 856–870, 2009.

[115] T. J. Drake, C. D. Medley, A. Sen, R. J. Rogers, and W. Tan, "Stochasticity of manganese superoxide dismutase mRNA expression in breast carcinoma cells by molecular beacon imaging," *ChemBioChem*, vol. 6, no. 11, pp. 2041–2047, 2005.

[116] H. Y. Yeh, M. V. Yates, A. Mulchandani, and W. Chen, "Visualizing the dynamics of viral replication in living cells via Tat peptide delivery of nuclease-resistant molecular beacons," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 45, pp. 17522–17525, 2008.

SAGE-Hindawi Access to Research
Journal of Nucleic Acids
Volume 2011, Article ID 102431, 9 pages
doi:10.4061/2011/102431

Review Article

Resveratrol, MicroRNAs, Inflammation, and Cancer

Esmerina Tili[1] and Jean-Jacques Michaille[1,2]

[1] *Department of Molecular Virology, Immunology, and Medical Genetics, Ohio State University, Biomedical Research Tower, 460 W 12th Avenue, Columbus, OH 43210, USA*
[2] *LBMN-INSERM U866, Université de Bourgogne, Faculté Gabriel, 6 boulevard Gabriel, 21000 Dijon, France*

Correspondence should be addressed to Esmerina Tili, tili.1@osu.edu and
Jean-Jacques Michaille, jean-jacques.michaille@u-bourgogne.fr

Received 20 October 2010; Revised 15 June 2011; Accepted 22 June 2011

Academic Editor: Dmitry A. Stetsenko

MicroRNAs are short noncoding RNAs that regulate the expression of many target genes posttranscriptionally and are thus implicated in a wide array of cellular and developmental processes. The expression of *miR-155* or *miR-21* is upregulated during the course of the inflammatory response, but these microRNAs are also considered oncogenes due to their upregulation of expression in several types of tumors. Furthermore, it is now well established that inflammation is associated with the induction or the aggravation of nearly 25% of cancers. Therefore, the above microRNAs are thought to link inflammation and cancer. Recently, resveratrol (trans-3,4′,5-trihydroxystilbene), a natural polyphenol with antioxidant, anti-inflammatory, and anticancer properties, currently at the stage of preclinical studies for human cancer prevention, has been shown to induce the expression of *miR-663*, a tumor-suppressor and anti-inflammatory microRNA, while downregulating *miR-155* and *miR-21*. In this paper we will discuss how the use of resveratrol in therapeutics may benefit from the preanalyses on the status of expression of *miR-155* or *miR-21* as well as of *TGFβ1*. In addition, we will discuss how resveratrol activity might possibly be enhanced by simultaneously manipulating the levels of its key target microRNAs, such as *miR-663*.

1. Inflammation and Cancer

Inflammation represents a complex, nonspecific immune response of the body to pathogens, damaged cells, tissue injury, allergens, toxic compounds, or irritant molecules [1]. While it is normally self-contained, it may become permanent and chronic. It may also escape the original tissue and spread via the circulatory and/or the lymphatic system to other parts of the body, producing a systemic inflammatory response syndrome such as sepsis in case of an infection. Chronic inflammation is associated with a simultaneous destruction and healing of the tissue from the inflammatory process and has been linked to a number of pathologies, including cancer, chronic asthma, rheumatoid arthritis, multiple sclerosis, inflammatory bowel diseases, and psoriasis, as well as several types of neurological disorders. The inflammatory response is coordinated by a large range of mediators that form complex regulatory networks [1]. The recruitment of leukocytes in the peripheral tissues is the hallmark of inflammation. It is mediated by several types of chemokines [1–3], which act through their receptors located at the surface of the cytoplasmic membrane of leukocytes. The production of chemokines is induced by inflammatory stimuli such as bacterial lipopolysaccharide (LPS), interleukin (IL)-1, or tumor necrosis factor (TNF). In addition, these chemokines have a clear role in angiogenesis and wound repair [4]. Finely tuned molecular mechanisms exist that ensure that the immune response required for the defense of the body has a limited duration and does not pass a certain maximum activation level, thus avoiding that, it becomes harmful to the organism. It is well acknowledged that unresolved immune response results in inflammation.

Epidemiological studies suggest that as many as 25% of all cancers may be due to chronic inflammation [5–7]. The connection between inflammation and cancer consists of

an extrinsic pathway, driven by inflammatory conditions that increase cancer risk, and an intrinsic pathway, driven by genetic alterations that cause inflammation and neoplasia [6]. Inflammatory mediators released by cancer-related inflammation induce genetic instability, leading to the accumulation of random genetic alterations in cancer cells [7]. The activation of Toll-like receptors (TLRs), a group of pattern recognition receptors functioning as sensors of pathogens and tissue damages, leads to the nuclear translocation of NF-κB and the production of cytokines such as TNF, IL-1α, IL-1β, IL-6, and IL-8. However, TLR activation has been shown to accelerate the growth of adoptively transferred tumors [8–11]. Accordingly, the stimulation of TLRs leads to increased survival and proliferation of several cell lines [12, 13], and the intratumoral injection of *Listeria monocytogenes* induces TLR2 signaling in tumor cells, thus promoting their growth [14]. TLR signaling also enhances tumor cell invasion and metastasis by regulating metalloproteinases and integrins [15]. Chemokines also affect several tumor progression pathways, such as leukocyte recruitment and function, cellular proliferation, survival, or senescence, as well as invasion and metastasis and are the targets of a number of anticancer agents [5].

Tumor microenvironment contains various inflammatory cell types infiltrating the tumor area in response to inflammatory stimuli, such as macrophages, neutrophils, and mast cells [16, 17]. Tumor-associated macrophages (TAMs) are thought to play key roles in the production of various growth factors, angiogenic factors, proteinases, chemokines, and cytokines, through crosstalks with cancer cells and other tumor stromal cells [18–20]. Factors secreted by TAMs stimulate cell migration/motility, proliferation, survival, angiogenesis, and metastasis, resulting in a dynamic environment that favors the progression of cancer, thus affecting the clinical outcome of malignant tumors. TAMs have thus been described as "obligate partners for tumor-cell migration, invasion and metastasis" [21, 22]. Namely, in a genetic model of breast cancer in macrophage-deficient mice, the tumors developed normally but were unable to form pulmonary metastases in the absence of macrophages [23]. As tumor metastasis is responsible for approximately 90% of all cancer-related deaths, a better understanding of inflammation regulatory mechanisms may potentially allow to optimize the use of anticancer drugs that lower tumor-specific inflammatory response [18].

Finally, the transforming growth factor β (TGFβ) regulates the immune response as well as the effects of the immune system on tumor progression or regression *in vivo* [24]. TGFβ has been shown to suppress the antitumor activity of T cells, natural killer (NK) cells, neutrophils, monocytes, and macrophages, which together are able to promote or repress tumor progression depending on the cellular context [25–27]. Importantly, TGFβ1, the most abundant and ubiquitously expressed isoform of TGFβ, is usually considered a tumor-suppressor, due to its cytostatic activity in epithelia. However, on advanced stages of tumors, TGFβ1 behaves as a tumor promoter, due to its capability to enhance angiogenesis, epithelial-to-mesenchymal transition, cell motility, and metastasis [28–30].

2. MicroRNAs and Inflammation

MicroRNAs (miRNAs) are short noncoding RNAs which regulate the translation and/or degradation of target messenger RNAs [31–33]. They have been implicated in the regulation of a number of fundamental processes, including muscle, cardiac, neural, and lymphocyte development, or the regulation of both the innate and adaptative immune responses [34, 35]. miRNAs originate from primary transcripts (pri-miRNAs) converted in the nucleus into precursor miRNAs (pre-miRNAs) by the RNase III Drosha, associated with DGCR8 to form the small microprocessor complex [36]. Pre-miRNAs are then exported in the cytoplasm where the miRNA hairpin is cleaved by the RNase III Dicer within the RISC loading complex. The guide strand, which corresponds to the mature miRNA, is then incorporated into the RISC complex [36]. miRNAs and their transcriptional regulators usually form autoregulatory loops aimed at controlling their respective levels [37]. miRNAs participate in many gene regulatory networks whose molecular malfunctions are associated with major pathologies such as cancer [34] or auto-mmune diseases [38–40].

Several miRNAs have been implicated in both inflammation and cancer [38–41]. The most prominent are *miR-155*, *miR-21,* and *miR-125b*. Thus the expression of *miR-155* is strongly elevated in several human leukemias and lymphomas ([40] and references therein). Transgenic mice with B cells overexpressing *miR-155* develop B-cell leukemia, and a sustained expression of *miR-155* in hematopoietic stem cells causes a myeloproliferative disorder [40]. On the other side, *miR-155* has been implicated in the regulation of myelopoiesis and erythropoiesis, Th1 differentiation, B-cell maturation, IgG1 production, somatic hypermutations, gene conversion, class switch recombination, and B- and T-cell homeostasis as well as in the regulation of the innate immune response [40]. Thus, *miR-155* levels increase following LPS treatment of Raw-264 macrophages, and *miR-155* transgenic mice show enhanced sensibility to LPS-induced endotoxin shock [42]. In contrast, *miR-155*-knock-out mice are unable to mount a proper T-cell or B-cell immune response [40]. The expression of another miRNA, *miR-125b*, was repressed within 1 hour of LPS challenge in Raw-264 cells [42]. In topic eczema, *miR-125b* expression was reduced in regions of the skin that were inflamed, while that of *miR-21* was enhanced [43]. Similarly, *miR-21* expression was increased by inflammation due to ulcerative colitis [44] and also by IL-13 and by specific antigens in OVA- and *Aspergillus fumigatus* antigens-induced asthma models [45]. The expression of *miR-21* changes dynamically during antigen-induced T-cell differentiation, with the highest levels of expression in the effector T cell [46]. *MiR-21* induction upon T-cell receptor (TCR) stimulation is believed to be involved in a negative feedback loop regulating TCR signaling [46]. *MiR-663* has drown recent attention due to its role not only as an anti-inflammatory miRNA but also as a tumor suppressor miRNA. Thus, *MiR-663* impairs the upregulation of *miR-155* by inflammatory stimuli [47]. In addition, the expression of this microRNA is lost in certain cancers such as gastric or pancreatic cancer, and it induces mitotic catastrophe

growth arrest when its expression is restored in these cells [48].

3. MicroRNAs as Oncogenes and Tumor-Suppressor Genes

miRNAs participate in many gene regulatory networks whose molecular malfunctions are associated with cancers [35, 41]. Depending on the effects of their downregulation or overexpression, miRNAs have been described either as oncogenic (onco-miRs) or tumor suppressors. Thus, the *miR-17-92* cluster on chromosome 13, which contains six miRNAs (*miR-17*, *-18a*, *-19a*, *-20a*, *-19b-1,* and *-92a-1*), is amplified and overexpressed in B-cell lymphomas and solid tumors such as breast or small-cell lung cancers, where it may enhance oncogenesis by potentially targeting E2F1, p21/CDKN1A, and BCL2L11/BIM [49]. On the other hand, loss of function of *miR-17-92* miRNAs might be advantageous for cancer cells in certain settings. Namely, loss of heterozygosity at the 13q31.3 locus has been observed in multiple tumor types, and a genome-wide analysis of copy number alterations in cancer revealed that the *miR-17-92* cluster was deleted in 16.5% of ovarian cancers, 21.9% of breast cancers, and 20% of melanomas [50]. In contrast, the twelve members of the human *let-7* gene family are frequently downregulated in cancers like lung, colon, or other solid tumors [34, 35]s and are, therefore,s considered as tumor-suppressor miRNAs in these types of cancers. In particular, *let-7* miRNAs target oncogenes of the *Ras* family [51] and c-Myc, and their expression in colon tumors results in reduced levels of both Ras and c-Myc [52]. *miR-21* is overexpressed in several cancers, including colorectal carcinomas (CRCs), gliomas, as well as breast, gastric, prostate, pancreas, lung, thyroid, and cervical cancers [53, 54]. *miR-21* has been shown to function as an onco-miR, due to its targeting of transcripts encoding key regulators of cell proliferation and apoptosis such as PTEN and PDCD4 [53]. Beside *miR-21*, several miRNAs are overexpressed in CRCs, including *miR-17*, *miR-25*, *miR-26a,* and *miR-181a* [54, 55].

In human, the levels of both *miR-155* and *BIC* transcripts (that is, *miR-155* primary RNAs) are elevated in diffuse large B-cell lymphoma (DLBCL), Hodgkins lymphoma, and primary mediastinal B-cell lymphoma [40]. In contrast, a very weak expression of *miR-155* is found in most non-Hodgkins lymphoma subtypes, including Burkitt lymphoma. In addition, high levels of *BIC* and *miR-155* expression were reported in B-cell chronic lymphocytic leukemia (B-CLL) and in B-CLL proliferation centers [56]. Furthermore, *miR-155* was also upregulated in a subset of patients with acute myelomonocytic leukemia and acute monocytic leukemia. Accordingly, transgenic mice whose B cells overexpress *miR-155* developed polyclonal preleukemic pre-B-cell proliferation followed by B-cell malignancy [40]. On the other hand, it was reported that *BIC* cooperates with *c-Myc* in avian lymphomagenesis and erythroleukemogenesis [57]. Beside liquid malignancies, high levels of *miR-155* expression were found in solid tumors such as breast, colon, and lung cancers [40]. *miR-155* was recently shown to induce a mutator phenotype by targeting Wee-1, a kinase regulating G_2/M phase transition during the cell cycle [58]. Furthermore, it was shown that *miR-155* increases genomic instability by targeting transcripts encoding components of the DNA mismatch repair machinery [59]. As a consequence, the simultaneous *miR-155*-steered suppression of a number of tumor suppressor genes combined with a mutator phenotype might allow the shortening of steps required for tumorogenesis, and might also explain how chronic inflammation associated with high levels of *miR-155* induces cancer.

Finally, miRNAs have been implicated in metastasis. For example, it has been established that the downregulation of both *miR-103-1* and *miR-103-2* miRNAs induces epithelial-to-mesenchymal transition by targeting *Dicer1* transcripts [60]. Furthermore, several miRNAs, including *miR-21*, have been shown to activate metastasis by acting on multiple signaling pathways and targeting various proteins that are involved in this process. Thus, in breast cancer, which represents the most common malignancy among women in the world, miRNAs such as *miR-9*, *miR-10b*, *miR-21*, *miR-103/107*, *miR-132*, *miR-373,* and *miR-520* stimulate metastasis, while *miR-7*, *miR-30*, *miR-31*, *miR-126*, *miR-145*, *miR-146*, *miR-200*, *miR-205*, *miR-335*, *miR-661,* and miRNAs of the *let-7* families in contrast impair the different steps of metastatic process, from epithelial-to-mesenchymal transition to local invasion to colonisation and angiogenesis [61].

Numerous reports have provided strong evidence that all of the above miRNAs potentially target a myriad of transcripts including those encoding transcription factors, cytokines, enzymes and kinases, implicated in both cancer and inflammation.

4. Anti-Inflammatory and Antitumor Properties of Resveratrol

Resveratrol (trans-3,4′,5-trihydroxystilbene) is a natural polyphenolic, nonflavonoid antioxidant found in grapes and other berries, produced by plants in response to infection by the pathogen *Botrytis cinerea* [62]. Recent studies have documented that resveratrol has various health benefits, such as cardiovascular- and cancer-preventive properties [63–65], and this compound is currently at the stage of preclinical studies for human cancer prevention [66, 67]. Resveratrol was first shown to inhibit both tumor promotion and tumor progression in a mouse skin cancer model [68]. Resveratrol is also tested for preventing and/or treating obesity and diabetes [69, 70]. Fortunately, resveratrol toxicity is minimal, and even proliferating tissues such as bone marrow or intestinal tract are not adversely affected [71].

Resveratrol exerts its effects at multiple levels. Both its *m*-hydroxyquinone and 4-hydroxystyryl moieties have been shown to be important for the determination of resveratrol inhibitory properties toward various enzymes. This include lipoxygenases and cyclooxygenases that synthesize proinflammatory mediators from arachidonic acid, protein kinases such as PKCs and PKD, and receptor tyrosine kinases, lipid kinases, as well as IKKα, an activator of NF-κB pathway, which establishes a strong link between inflammation

and tumorigenesis [72]. Also, resveratrol inhibition of P450/CYP19A1/Aromatase, by limiting the amount of available estrogens and consequently the activity of estrogen receptors, has been proposed to contribute to the protection against several types of cancer, including breast cancer [72, 73]. Of note, resveratrol also inhibits the formation of estrogen-DNA adducts, which are elevated in women at high risk for breast cancer [74].

Resveratrol in addition regulates apoptosis and cell proliferation. It induces growth arrest followed by apoptotic cell death and interferes with cell survival by upregulating the expression of proapoptotic genes while simultaneously downregulating the expression of antiapoptotic genes [75]. Resveratrol induces the redistribution of CD95 and other death receptors in lipid rafts, thus contributing to their sensitization to death receptor agonists [75]. It also causes growth arrest at G1 and G1/S phases of cell cycle by inducing the expression of CDK inhibitors p21/CDKN1A and p27/CDKN1B [63]. In addition, resveratrol directly inhibits DNA synthesis by diminishing ribonucleotide reductase and DNA polymerase [72, 76, 77]. Altogether, antiproliferative activities of resveratrol involve the differential regulation of multiple cell-cycle targets in a cell-type-dependent manner [72, 75].

One of the possible mechanisms for resveratrol protective activities is by downregulation of the inflammatory responses [78]. That includes the inhibition of synthesis and release of proinflammatory mediators, modifications of eicosanoid synthesis, or inhibiting the enzymes, such as cyclooxygenase-1 (COX-1/PTGS1) or -2 (COX-2/PTGS2), which are responsible for the synthesis of proinflammatory mediators, through the inhibitory effect of resveratrol on transcription factors like NF-κB or activator protein-1 (AP-1) [78, 79]. Of note, constitutive COX-2 expression generally predicts aggressiveness of tumors, therefore, the use of nonsteroidal anti-inflammatory drugs that inhibit COX-2 in cancer treatment. However, cytoplasmic COX-2 can relocalize in the nucleus. This nuclear relocalization of COX-2 is induced by resveratrol, and exposure of resveratrol-treated cells to a specific COX-2 inhibitor blocked resveratrol-induced apoptosis, indicating that COX-2 displays proapoptotic activity in the nucleus, which may be associated with the generation of complexes of COX-2 and ERK1/2 mitogen-activated protein kinases. In mouse macrophages, resveratrol also displays antioxidant activity, decreasing the production of reactive oxygen species and reactive nitrogen species and inhibiting nitric oxygen synthetase (NOS)-2 and COX-2 synthesis as well as prostaglandin E2 production [80].

Furthermore, in 3T3-L1 adipocytes, resveratrol inhibits the production of the TNF-induced monocyte chemoattractant protein (MCP)-1/CCL2. MCP-1 plays an essential role in the early events during macrophage infiltration into adipose-tissue, which results in chronic low-grade inflammation, a key feature of obesity type 2 diabetes characterized by adipose tissue macrophage infiltration and abnormal cytokine production [81]. Finally, it is also well established that some anti-inflammatory effects of resveratrol arise from its capability to upregulate histone deacetylase sirtuin 1 (SIRT1) activity, that also presents antitumor and anti-inflammatory capabilities [82]. Altogether, it is clear that the key antitumor properties of resveratrol [77–82] are linked to its anti-inflammatory effects.

The fact that resveratrol targets, directly or indirectly, so many different factors, exerts such a wide influence on cell homeostasis, and provides such a range of health benefits, suggested that some of its effects should arise from its capability to modulate the activity of global regulators. Furthermore, the ability of each miRNA to potentially regulate the levels and, therefore, the activity, of tens to hundreds of target genes, strongly suggested that resveratrol should be able to modify the composition of miRNA populations. Indeed, it has recently been shown that resveratrol decreases the levels of several proinflammatory and/or oncogenic miRNAs and upregulates miRNAs with anti-inflammatory and/or antitumor potentials [47, 83].

5. *MiR-663* as a Mediator of Resveratrol Anti-Inflammatory Activity

Affymetrix microarrays and RNase-protection assays showed that resveratrol treatment of human THP-1 monocytic cells upregulated the expression of *LOC284801* transcripts, that contain the sequence of *pre-miR-663* and thus represent *miR-663* primary transcripts. MiRNA microarrays and RNase-protection assays accordingly confirmed these data [47]. Interestingly, *in silico* analysis using TargetScan (http://www.targetscan.org/) suggested that *miR-663* may potentially target transcripts encoding factors implicated in (i) the mounting of the immune response, especially JunB, JunD, and FosB, which encode AP-1 factors known to activate many cytokine genes in partnership with NFAT factors [84], (ii) TLR signaling, such as the kinases RIPK1 and IRAK2, and (iii) the differentiation of monocytes, Th1 lymphocytes, and granulocytes.

An antisense *miR-663* inhibitory RNA (*663-I*) proved capable of increasing global AP-1 activity in unchallenged THP-1 cells, showing that *miR-663* indeed target transcripts encoding AP-1 factors in these cells. These effects were in particular directed toward *JunB* and *JunD* transcripts [47]. In agreement with previous results [79], resveratrol blocked the surge of AP-1 activity that occurs following LPS challenge due to the fact that *JunB* transcripts peak within the first hour, leading to the accumulation of JunB in the next few hours [85]. This inhibitory effect of resveratrol on AP-1 activity was partly impaired by *663-I*, indicating that it arises at least in part from the upregulation of *miR-663* by resveratrol [47]. Western blots showed that resveratrol impaired JunB neosynthesis, while still allowing the phosphorylation, that is, the activation of JunB following LPS treatment to take place, at least to a certain extent. Given that AP-1 factors include c-Jun, JunB, JunD, FosB, Fra-1, and Fra-2, as well as Jun dimerization partners JDP1 and JDP2 or the closely related ATF2, LRF1/ATF3, and B-ATF, so that potentially about 18 different dimeric combinations may be formed, the capability of resveratrol to specifically target a subset of AP-1 dimers through the upregulation of *miR-663* might have profound effects on the levels of

the transcriptional activity of promoters to whom different AP-1 factors can compete for binding. Due to the many roles of AP-1 factors both in inflammation and cancer [86, 87], the specific targeting of genes encoding a subset of AP-1 factors, by changing the composition of AP-1 dimers on key promoters, may possibly explain some of the multiple anti-inflammatory and anticancer properties of resveratrol.

Of note, *miR-155* had been shown to be under AP-1 activity in activated B cells [88]. Accordingly, *miR-663* reduced the upregulation of *miR-155* by LPS [42], which may be due to *miR-663* targeting of transcripts encoding JunB and JunD and also possibly FosB and KSRP, an RNA binding protein implicated in the LPS-induced *miR-155* maturation from its primary transcripts *BIC* [89]. This is of primary importance, for *miR-155* upregulation is the hallmark of inflammatory response following LPS treatment of macrophages/monocytes [42]. Resveratrol also dramatically impaired the upregulation of *miR-155* by LPS, an effect partly inhibited by *663-I* [47]. Altogether, these results indicate that the anti-inflammatory properties of resveratrol arise, at least in part, from its upregulation of *miR-663* and its downregulating effects on *miR-155* and that *miR-663* might possibly qualify as an anti-inflammatory miRNA.

6. MicroRNAs as Mediators of Resveratrol Anticancer Effects

The results reported here above also suggested that, due to its targeting of AP-1 factors, known to play a role in tumorigenesis and cell invasion [86, 87] and due to its downregulation of *miR-155*, whose levels increase in solid as well as in liquid tumors [40], *miR-663* may also possibly provide resveratrol with some of its anticancer properties. Namely, *miR-663* was found to be downregulated in hormone refractory prostate cancer cells, along with *miR-146a* and *miR-146b* [90], further supporting the hypothesis that this miRNA is a tumor-suppressor gene whose one of the function is to keep low the expression level of oncogenic *miR-155* [47, 48].

CRC is the third most common malignancy and the fourth biggest cause of cancer mortality worldwide [91, 92]. Despite the increased use of screening strategies such as fecal occult blood testing, sigmoidoscopy, and colonoscopy, more than one-third of patients with colorectal cancer will ultimately develop metastatic disease [92]. On the other hand, the TGFβ signaling pathway is one of the most commonly altered cellular signaling pathways in human cancers [93]. Among the three TGFβ isoforms expressed in mammalian epithelia (TGFβ1, TGFβ2, and TGFβ3), TGFβ1 is the most abundant and ubiquitously expressed one. TGFβ signaling is initiated by the binding of TGFβ ligands to type II receptors (TGFβR2). Once bound by TGFβ, TGFβR2 recruits, phosphorylates, and thus activates the type I TGFβ receptor (TGFβR1). TGFβR1 then phosphorylates two transcriptional regulators, namely, SMAD2 and SMAD3, which subsequently bind to SMAD4. This results in the nuclear translocation of SMAD complexes, allowing SMADs to interact with transcription factors controlling the expression of a multitude of TGFβ responsive genes [94]. The expression of TGFβ1 in both tumor and plasma was found to be significantly higher in patients with metastatic colorectal cancer, and increasing colorectal tumor stage was correlated with higher TGFβ1 expression in tumor tissues [95].

miRNA microarrays recently showed that resveratrol treatment of SW480 human colon cancer cells significantly increased the levels of 22 miRNAs while decreasing those of 26 others [83]. Among the miRNAs downregulated by resveratrol, *miR-17*, *miR-21*, *miR-25*, *miR-92a-2*, *miR-103-1* and *miR-103-2* have been shown to behave as onco-miRNAs, at least in certain contexts. Thus, genomic amplification and overexpression of *miR-17-92* miRNAs is found in B-cell lymphomas as well as in breast and lung cancers [34, 54]. *MiR-21* is overexpressed in several cancers, including CRCs, gliomas, as well as breast, gastric, prostate, pancreas, lung, thyroid, and cervical cancers [53–55]. *MiR-17*, *miR-25*, *miR-26-a,* and *miR-181a* are also overexpressed in CRCs [34, 54]. In addition, several miRNAs, including *miR-21*, have been shown to activate metastasis by acting on multiple signaling pathways and targeting various proteins that are key players in this process [53]. Furthermore, the lower metastatic propensity of SW480 cells as compared with SW620 human colon cancer cells, both derived from the primary tumor and a metastasis of the same patient, respectively [96], was associated with a lower level of expression of *miR-103-1* and *miR-103-2*, two miRNAs that induce epithelial-to-mesenchymal transition by targeting *Dicer1* transcripts [60].

In silico analysis using TargetScan showed that miRNAs downregulated by resveratrol in SW480 cells potentially target transcripts encoding known tumor suppressor factors, such as the two antiproliferation factors PDCD4 and PTEN, the components of the mismatch repair machinery MLH3, MSH2 and MSH3, DICER1, the RNase III producing mature miRNAs from their immediate precursors in the cytoplasm, and several effectors and regulators of the TGFβ signaling pathway [83]. Indeed, resveratrol treatment of SW480 cells lead to a greater accumulation of TGFβR1, TGFβR2, PDCD4, PTEN, and E-CADHERIN (a component of adherens junctions implicated in the maintenance of epithelial phenotype) [83]. Of note, among miRNAs upregulated by resveratrol, *miR-663* was the only one to target *TGFβ1* transcripts. Luciferase assays and Western blots showed that resveratrol downregulated *TGFβ1* in both a *miR-663*-dependent and a *miR-663*-independent manner [83]. Resveratrol treatment also decreased the transcriptional activity of SMADs under *TGFβ1* signaling, an effect seemingly independent of *miR-663* [83].

Interestingly, it has been recently shown that GAM/ZNF512B, a vertebrate-specific developmental regulator first described in chicken [97], impairs the upregulation of miRNAs of the *miR-17-92* cluster by TGFβ1 and that TGFβ1 in turn downregulates *GAM*, at least in part through the upregulation of *miR-17-92* miRNAs [98]. The facts that *GAM* transcripts contain three consensus target sites for *miR-663* and that GAM is sensible to resveratrol treatment (Tili et al., unpublished results) raises the question of the possible existence of a gene regulatory network that would allow *miR-663* to impair GAM repressing activity on TGFβ1 signaling

pathway when TGFβ1 works as a tumor suppressor, that is, at the early stages of tumorigenesis but not any more when this pathway starts to favor tumorigenesis and metastasis, that is, on advanced stages of cancers.

It is important to emphasize that TGFβ1 has been shown to enhance the maturation of oncogenic *miR-21* through the binding of SMAD3 to *miR-21* primary RNAs [99] and also to increase the expression of *miR-155*, which is under the control of TGFβ/SMAD activity [100]. Thus, by targeting AP-1 factors as well as TGFβ1 and possibly SMAD3, *miR-663* might inhibit two of the pathways that upregulate *miR-155* expression.

The TGFβ signaling pathway present multiple levels of regulation: sequestration of ligands into inactive precursor forms, ligand traps, decoy receptors, and inhibitory SMADs, not to mention the existence of SMAD-independent pathways and their interactions with many other critical signaling pathways which also proved to play a role in cancer. It is thus not surprising that the short (307-nt long) 3′-UTR of *TGFβ1* transcripts contains a potential consensus target site for 28 miRNAs only. Of note, *TGFβ1* 3′-UTR contains two target sites for two of these miRNAs and only one target site for 25 of the others. Therefore, the fact that *miR-663* may potentially target 5 different sites in *TGFβ1* 3′-UTR suggests that this miRNA could represent a critical *TGFβ1* regulator which may possibly be called upon action in emergency situations such as those when cells begin to proliferate anarchically or when a stronger immune response is required. The multiplicity of *miR-663* targets sites in *TGFβ1* 3′-UTR further suggests that the effects of this miRNA might be both dose- and context-dependent, so that resveratrol effects on TGFβ1 signaling pathway might well be also context-dependent. Finally, it is probable that, depending on the cell context, resveratrol might either increase the level of TGFβ signaling—by inhibiting miRNAs targeting its main effectors—when it is beneficial to the organism, that is, when it works to maintain the integrity of epithelia and impair cell proliferation, or in contrast decrease TGFβ1 signaling—by decreasing its production through the upregulation of *miR-663*—when TGFβ1 starts to favor epithelial-to-mesenchymal transition and metastasis. For example, the targeting of both *TGFβ1* and *SMAD3* transcripts might possibly allow resveratrol to impair TGFβ1-induced SMAD3-dependent promotion of cell motility and invasiveness in advanced stages of gastric cancer [101, 102] or when SMAD2 and SMAD3 phosphorylated at both linker and COOH-terminal regions transmit malignant TGFβ1 signal in later stages of human CRC [103].

7. Conclusions

It is notable that striking phenotypes are often driven through small changes in the cellular concentration of key factors. For example, in the B-cell compartment, *miR-150* curtails the activity of c-Myb transcription factor in a dose-dependent fashion over a narrow range of miRNA and c-Myb concentrations [104]. Thus, even slight effects of resveratrol on a handful of key miRNAs might well prove to be critical to its anti-inflammatory, anticancer and antimetastatic properties. In addition, the fact that *miR-663*, *miR-21*, *miR-155,* and TGFβ1 have all been implicated in the regulation of cell proliferation, tumor apparition and development, metastasis formation, and innate immunity, strongly suggests that the capability of resveratrol to behave at the same time as an antitumor, antimetastatic, antiproliferation and anti-inflammatory agent most probably arises from its effects on the expression of a small set of critical endogenous miRNAs having the abilities to impact the cell proteome globally.

Finally, miRNAs have the promise to become biomarkers for different stages of cancer, both for diagnosis and prognosis. Furthermore, the discovery that resveratrol can modulate the levels of miRNAs targeting proinflammatory and/or protumor factors opens the possibility to optimize resveratrol treatments by manipulating in parallel the levels of expression of a few critical miRNAs. For example, from the experiments reported here above, it starts to become clear that the use of resveratrol would be especially beneficiary in the type of cancers where the TGFβ pathway is implicated. Of course, resveratrol use would have to be carefully correlated with the stages of cancers, knowing that TGFβ can have two faces, that is, anti- and prometastatic.

As a last remark, it should be noted that while resveratrol antitumor potential has been linked with data primarily from human cell culture systems, evidence that resveratrol can inhibit carcinogenesis in several organ sites emerged from results of cancer prevention and therapy studies in laboratory animal models [68]. Given that *miR-663* was only found in primates, the reports by Tili et al. [47, 83] come as a warning that studies in animal may not always allow to predict accurately the molecular effects of resveratrol in human, especially when it comes to miRNAs.

References

[1] R. Medzhitov, "Origin and physiological roles of inflammation," *Nature*, vol. 454, no. 7203, pp. 428–435, 2008.

[2] C. Gerard and B. J. Rollins, "Chemokines and disease," *Nature Immunology*, vol. 2, no. 2, pp. 108–115, 2001.

[3] I. F. Charo and R. M. Ransohoff, "The many roles of chemokines and chemokine receptors in inflammation," *The New England Journal of Medicine*, vol. 354, no. 6, pp. 610–621, 2006.

[4] A. Mantovani, B. Savino, M. Locati, L. Zammataro, P. Allavena, and R. Bonecchi, "The chemokine system in cancer biology and therapy," *Cytokine and Growth Factor Reviews*, vol. 21, no. 1, pp. 27–39, 2010.

[5] S. P. Hussain and C. C. Harris, "Inflammation and cancer: an ancient link with novel potentials," *International Journal of Cancer*, vol. 121, no. 11, pp. 2373–2380, 2007.

[6] A. Mantovani, P. Allavena, A. Sica, and F. Balkwill, "Cancer-related inflammation," *Nature*, vol. 454, no. 7203, pp. 436–444, 2008.

[7] F. Colotta, P. Allavena, A. Sica, C. Garlanda, and A. Mantovani, "Cancer-related inflammation, the seventh hallmark of cancer: links to genetic instability," *Carcinogenesis*, vol. 30, no. 7, pp. 1073–1082, 2009.

[8] S. Rakoff-Nahoum and R. Medzhitov, "Toll-like receptors and cancer," *Nature Reviews Cancer*, vol. 9, no. 1, pp. 57–63, 2009.

[9] L. Sfondrini, A. Rossini, D. Besusso et al., "Antitumor activity of the TLR-5 ligand flagellin in mouse models of cancer," *Journal of Immunology*, vol. 176, no. 11, pp. 6624–6630, 2006.

[10] G. P. Pidgeon, J. H. Harmey, E. Kay, M. da Costa, H. P. Redmond, and D. J. Bouchier-Hayes, "The role of endotoxin/lipopolysaccharide in surgically induced tumour growth in a murine model of metastatic disease," *British Journal of Cancer*, vol. 81, no. 8, pp. 1311–1317, 1999.

[11] J. H. Harmey, C. D. Bucana, W. Lu et al., "Lipopolysaccharide-induced metastatic growth is associated with increased angiogenesis, vascular permeability and tumor cell invasion," *International Journal of Cancer*, vol. 101, no. 5, pp. 415–422, 2002.

[12] G. Jego, R. Bataille, A. Geffroy-Luseau, G. Descamps, and C. Pellat-Deceunynck, "Pathogen-associated molecular patterns are growth and survival factors for human myeloma cells through Toll-like receptors," *Leukemia*, vol. 20, no. 6, pp. 1130–1137, 2006.

[13] J. Bohnhorst, T. Rasmussen, S. H. Moen et al., "Toll-like receptors mediate proliferation and survival of multiple myeloma cells," *Leukemia*, vol. 20, no. 6, pp. 1138–1144, 2006.

[14] B. Huang, J. Zhao, S. Shen et al., "Listeria monocytogenes promotes tumor growth via tumor cell toll-like receptor 2 signaling," *Cancer Research*, vol. 67, no. 9, pp. 4346–4352, 2007.

[15] B. Huang, J. Zhao, H. Li et al., "Toll-like receptors on tumor cells facilitate evasion of immune surveillance," *Cancer Research*, vol. 65, no. 12, pp. 5009–5014, 2005.

[16] F. Balkwill and A. Mantovani, "Inflammation and cancer: back to Virchow?" *The Lancet*, vol. 357, no. 9255, pp. 539–545, 2001.

[17] L. M. Coussens and Z. Werb, "Inflammation and cancer," *Nature*, vol. 420, no. 6917, pp. 860–867, 2002.

[18] M. Ono, "Molecular links between tumor angiogenesis and inflammation: inflammatory stimuli of macrophages and cancer cells as targets for therapeutic strategy," *Cancer Science*, vol. 99, no. 8, pp. 1501–1506, 2008.

[19] J. W. Pollard, "Tumour-educated macrophages promote tumour progression and metastasis," *Nature Reviews Cancer*, vol. 4, no. 1, pp. 71–78, 2004.

[20] A. Yoshimura, "Signal transduction of inflammatory cytokines and tumor development," *Cancer Science*, vol. 97, no. 6, pp. 439–447, 2006.

[21] A. Sica and V. Bronte, "Altered macrophage differentiation and immune dysfunction in tumor development," *Journal of Clinical Investigation*, vol. 117, no. 5, pp. 1155–1165, 2007.

[22] J. Condeelis and J. W. Pollard, "Macrophages: obligate partners for tumor cell migration, invasion, and metastasis," *Cell*, vol. 124, no. 2, pp. 263–266, 2006.

[23] E. Y. Lin, A. V. Nguyen, R. G. Russell, and J. W. Pollard, "Colony-stimulating factor 1 promotes progression of mammary tumors to malignancy," *Journal of Experimental Medicine*, vol. 193, no. 6, pp. 727–740, 2001.

[24] B. Bierie and H. L. Moses, "Transforming growth factor beta (TGF-β) and inflammation in cancer," *Cytokine and Growth Factor Reviews*, vol. 21, no. 1, pp. 49–59, 2010.

[25] M. O. Li, Y. Y. Wan, S. Sanjabi, A. K. L. Robertson, and R. A. Flavell, "Transforming growth factor-β regulation of immune responses," *Annual Review of Immunology*, vol. 24, pp. 99–146, 2006.

[26] S. H. Wrzesinski, Y. Y. Wan, and R. A. Flavell, "Transforming growth factor-β and the immune response: implications for anticancer therapy," *Clinical Cancer Research*, vol. 13, no. 18, pp. 5262–5270, 2007.

[27] J. Massagué, S. W. Blain, and R. S. Lo, "TGFβ signaling in growth control, cancer, and heritable disorders," *Cell*, vol. 103, no. 2, pp. 295–309, 2000.

[28] J. Massagué, "TGFβ in cancer," *Cell*, vol. 134, no. 2, pp. 215–230, 2008.

[29] D. Padua and J. Massagué, "Roles of TGFβ in metastasis," *Cell Research*, vol. 19, no. 1, pp. 89–102, 2009.

[30] M. Tian and W. P. Schiemann, "The TGFβ paradox in human cancer: an update," *Future Oncology*, vol. 5, no. 2, pp. 259–271, 2009.

[31] V. Ambros, "The functions of animal microRNAs," *Nature*, vol. 431, no. 7006, pp. 350–355, 2004.

[32] D. P. Bartel and C. Z. Chen, "Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs," *Nature Reviews Genetics*, vol. 5, no. 5, pp. 396–400, 2004.

[33] G. Stefani and F. J. Slack, "Small non-coding RNAs in animal development," *Nature Reviews Molecular Cell Biology*, vol. 9, no. 3, pp. 219–230, 2008.

[34] E. Tili, J. J. Michaille, V. Gandhi, W. Plunkett, D. Sampath, and G. A. Calin, "MiRNAs and their potential for use against cancer and other diseases," *Future Oncology*, vol. 3, no. 5, pp. 521–537, 2007.

[35] E. Tili, J. J. Michaille, S. Costinean, and C. M. Croce, "MicroRNAs, the immune system and rheumatic disease," *Nature Clinical Practice Rheumatology*, vol. 4, no. 10, pp. 534–541, 2008.

[36] J. Winter, S. Jung, S. Keller, R. I. Gregory, and S. Diederichs, "Many roads to maturity: microRNA biogenesis pathways and their regulation," *Nature Cell Biology*, vol. 11, no. 3, pp. 228–234, 2009.

[37] J. Tsang, J. Zhu, and A. van Oudenaarden, "MicroRNA-mediated feedback and feedforward loops are recurrent network motifs in mammals," *Molecular Cell*, vol. 26, no. 5, pp. 753–767, 2007.

[38] E. Sonkoly and A. Pivarcsi, "Advances in microRNAs: implications for immunity and inflammatory diseases," *Journal of Cellular and Molecular Medicine*, vol. 13, no. 1, pp. 24–38, 2009.

[39] M. Iborra, F. Bernuzzi, P. Invernizzi, and S. Danese, "MicroRNAs in autoimmunity and inflammatory bowel disease: crucial regulators in immune response," *Autoimmunity Reviews*, 2010.

[40] E. Tili, C. M. Croce, and J. J. Michaille, "MiR-155: on the crosstalk between inflammation and cancer," *International Reviews of Immunology*, vol. 28, no. 5, pp. 264–284, 2009.

[41] E. Tili, J. J. Michaille, and G. A. Calin, "Expression and function of micro-RNAs in immune cells during normal or disease state," *International Journal of Medical Sciences*, vol. 5, no. 2, pp. 73–79, 2008.

[42] E. Tili, J. J. Michaille, A. Cimino et al., "Modulation of miR-155 and miR-125b levels following lipopolysaccharide/TNF-α stimulation and their possible roles in regulating the response to endotoxin shock," *Journal of Immunology*, vol. 179, no. 8, pp. 5082–5089, 2007.

[43] E. Sonkoly, T. Wei, P. C. Janson et al., "MicroRNAs: novel regulators involved in the pathogenesis of psoriasis?" *PLoS One*, vol. 2, no. 7, p. e610, 2007.

[44] F. Wu, M. Zikusoka, A. Trindade et al., "MicroRNAs are differentially expressed in ulcerative colitis and alter expression of macrophage inflammatory peptide-2α," *Gastroenterology*, vol. 135, no. 5, article e24, pp. 1624–1635, 2008.

[45] T. X. Lu, A. Munitz, and M. E. Rothenberg, "MicroRNA-21 is up-regulated in allergic airway inflammation and regulates IL-12p35 expression," *Journal of Immunology*, vol. 182, no. 8, pp. 4994–5002, 2009.

[46] H. Wu, J. R. Neilson, P. Kumar et al., "MiRNA profiling of naïve, effector and memory CD8 T cells," *PLoS One*, vol. 2, no. 10, Article ID e1020, 2007.

[47] E. Tili, J. J. Michaille, B. Adair et al., "Resveratrol decreases the levels of miR-155 by upregulating miR-663, a microRNA targeting JunB and JunD," *Carcinogenesis*, vol. 31, no. 9, pp. 1561–1566, 2010.

[48] J. Pan, H. Hu, Z. Zhou et al., "Tumor-suppressive mir-663 gene induces mitotic catastrophe growth arrest in human gastric cancer cells," *Oncology Reports*, vol. 24, no. 1, pp. 105–112, 2010.

[49] J. T. Mendell, "MiRiad roles for the miR-17-92 cluster in development and disease," *Cell*, vol. 133, no. 2, pp. 217–222, 2008.

[50] L. Zhang, J. Huang, N. Yang et al., "MicroRNAs exhibit high frequency genomic alterations in human cancer," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 24, pp. 9136–9141, 2006.

[51] S. M. Johnson, H. Grosshans, J. Shingara et al., "RAS is regulated by the let-7 microRNA family," *Cell*, vol. 120, no. 5, pp. 635–647, 2005.

[52] Y. Akao, Y. Nakagawa, and T. Naoe, "Let-7 microRNA functions as a potential growth suppressor in human colon cancer cells," *Biological and Pharmaceutical Bulletin*, vol. 29, no. 5, pp. 903–906, 2006.

[53] A. M. Krichevsky and G. Gabriely, "MiR-21: a small multifaceted RNA," *Journal of Cellular and Molecular Medicine*, vol. 13, no. 1, pp. 39–53, 2009.

[54] S. Volinia, G. A. Calin, C. G. Liu et al., "A microRNA expression signature of human solid tumors defines cancer gene targets," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 7, pp. 2257–2261, 2006.

[55] O. Slaby, M. Svoboda, P. Fabian et al., "Altered expression of miR-21, miR-31, miR-143 and miR-145 is related to clinicopathologic features of colorectal cancer," *Oncology*, vol. 72, no. 5-6, pp. 397–402, 2007.

[56] G. A. Calin, M. Ferracin, A. Cimmino et al., "A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia," *The New England Journal of Medicine*, vol. 353, no. 17, pp. 1793–1801, 2005.

[57] W. Tam, S. H. Hughes, W. S. Hayward, and P. Besmer, "Avian bic, a gene isolated from a common retroviral site in avian leukosis virus-induced lymphomas that encodes a noncoding RNA, cooperates with c-myc in lymphomagenesis and erythroleukemogenesis," *Journal of Virology*, vol. 76, no. 9, pp. 4275–4286, 2002.

[58] E. Tili, J. J. Michaille, D. Wernicke et al., "Mutator activity induced by microRNA-155 (miR-155) links inflammation and cancer," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 12, pp. 4908–4913, 2011.

[59] N. Valeri, P. Gasparini, M. Fabbri et al., "Modulation of mismatch repair and genomic stability by miR-155," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 15, pp. 6982–6987, 2011.

[60] G. Martello, A. Rosato, F. Ferrari et al., "A microRNA targeting dicer for metastasis control," *Cell*, vol. 141, no. 7, pp. 1195–1207, 2010.

[61] M. Shi, D. Liu, H. Duan, B. Shen, and N. Guo, "Metastasis-related miRNAs, active players in breast cancer invasion, and metastasis," *Cancer and Metastasis Reviews*, vol. 29, pp. 785–799, 2010.

[62] D. Delmas, A. Lançon, D. Colin, B. Jannin, and N. Latruffe, "Resveratrol as a chemopreventive agent: a promising molecule for fighting cancer," *Current Drug Targets*, vol. 7, no. 4, pp. 423–442, 2006.

[63] S. Shankar, G. Singh, and R. K. Srivastava, "Chemoprevention by resveratrol: molecular mechanisms and therapeutic potential," *Frontiers in Bioscience*, vol. 12, pp. 4839–4854, 2007.

[64] A. Csiszar, K. Smith, N. Labinskyy, Z. Orosz, A. Rivera, and Z. Ungvari, "Resveratrol attenuates TNF-α-induced activation of coronary arterial endothelial cells: role of NF-κB inhibition," *American Journal of Physiology—Heart and Circulatory Physiology*, vol. 291, no. 4, pp. H1694–H1699, 2006.

[65] M. Athar, J. H. Back, L. Kopelovich, D. R. Bickers, and A. L. Kim, "Multiple molecular targets of resveratrol: anti-carcinogenic mechanisms," *Archives of Biochemistry and Biophysics*, vol. 486, no. 2, pp. 95–102, 2009.

[66] A. Bishayee, "Cancer prevention and treatment with resveratrol: from rodent studies to clinical trials," *Cancer Prevention Research*, vol. 2, no. 5, pp. 409–418, 2009.

[67] M. Athar, J. H. Back, X. Tang et al., "Resveratrol: a review of preclinical studies for human cancer prevention," *Toxicology and Applied Pharmacology*, vol. 224, no. 3, pp. 274–283, 2007.

[68] M. Jang, L. Cai, G. O. Udeani et al., "Cancer chemopreventive activity of resveratrol, a natural product derived from grapes," *Science*, vol. 275, no. 5297, pp. 218–220, 1997.

[69] K. Szkudelska and T. Szkudelski, "Resveratrol, obesity and diabetes," *European Journal of Pharmacology*, vol. 635, no. 1–3, pp. 1–8, 2010.

[70] T. E. Sonnett, T. L. Levien, B. J. Gates, J. D. Robinson, and R. K. Campbell, "Diabetes mellitus, inflammation, obesity: proposed treatment pathways for current and future therapies," *Annals of Pharmacotherapy*, vol. 44, no. 4, pp. 701–711, 2010.

[71] J. A. Crowell, P. J. Korytko, R. L. Morrissey, T. D. Booth, and B. S. Levine, "Resveratrol-associated renal toxicity," *Toxicological Sciences*, vol. 82, no. 2, pp. 614–619, 2004.

[72] L. Pirola and S. Fröjdö, "Resveratrol: one molecule, many targets," *IUBMB Life*, vol. 60, no. 5, pp. 323–332, 2008.

[73] L. Le Corre, N. Chalabi, L. Delort, Y. J. Bignon, and D. J. Bernard-Gallon, "Resveratrol and breast cancer chemoprevention: molecular mechanisms," *Molecular Nutrition and Food Research*, vol. 49, no. 5, pp. 462–471, 2005.

[74] E. L. Cavalieri and E. G. Rogan, "Depurinating estrogen-DNA adducts in the etiology and prevention of breast and other human cancers," *Future Oncology*, vol. 6, no. 1, pp. 75–91, 2010.

[75] D. Delmas, C. Rébé, O. Micheau et al., "Redistribution of CD95, DR4 and DR5 in rafts accounts for the synergistic toxicity of resveratrol and death receptor ligands in colon carcinoma cells," *Oncogene*, vol. 23, no. 55, pp. 8979–8986, 2004.

[76] M. Fontecave, M. Lepoivre, E. Elleingand, C. Gerez, and O. Guittet, "Resveratrol, a remarkable inhibitor of ribonucleotide reductase," *FEBS Letters*, vol. 421, no. 3, pp. 277–279, 1998.

[77] N. J. Sun, S. H. Woo, J. M. Cassady, and R. M. Snapka, "DNA polymerase and topoisomerase II inhibitors from Psoralea corylifolia," *Journal of Natural Products*, vol. 61, no. 3, pp. 362–366, 1998.

[78] S. Das and D. K. Das, "Anti-inflammatory responses of resveratrol," *Inflammation and Allergy—Drug Targets*, vol. 6, no. 3, pp. 168–173, 2007.

[79] J. K. Kundu, Y. K. Shin, and Y. J. Surh, "Resveratrol modulates phorbol ester-induced pro-inflammatory signal transduction pathways in mouse skin in vivo: NF-κB and AP-1 as prime targets," *Biochemical Pharmacology*, vol. 72, no. 11, pp. 1506–1515, 2006.

[80] J. Leiro, E. Álvarez, J. A. Arranz, R. Laguna, E. Uriarte, and F. Orallo, "Effects of cis-resveratrol on inflammatory murine macrophages: antioxidant activity and down-regulation of inflammatory genes," *Journal of Leukocyte Biology*, vol. 75, no. 6, pp. 1156–1165, 2004.

[81] J. Zhu, W. Yong, X. Wu et al., "Anti-inflammatory effect of resveratrol on TNF-α-induced MCP-1 expression in adipocytes," *Biochemical and Biophysical Research Communications*, vol. 369, no. 2, pp. 471–477, 2008.

[82] S. Chung, H. Yao, S. Caito, J. W. Hwang, G. Arunachalam, and I. Rahman, "Regulation of SIRT1 in cellular functions: role of polyphenols," *Archives of Biochemistry and Biophysics*, vol. 501, no. 1, pp. 79–90, 2010.

[83] E. Tili, J. J. Michaille, H. Alder et al., "Resveratrol modulates the levels of microRNAs targeting genes encoding tumor-suppressors and effectors of TGFβ signaling pathway in SW480 cells," *Biochemical Pharmacology*, vol. 80, no. 12, pp. 2057–2065, 2010.

[84] F. Macián, C. López-Rodríguez, and A. Rao, "Partners in transcription: NFAT and AP-1," *Oncogene*, vol. 20, no. 19, pp. 2476–2489, 2001.

[85] M. Fujihara, M. Muroi, Y. Muroi, N. Ito, and T. Suzuki, "Mechanism of lipopolysaccharide-triggered junB activation in a mouse macrophage-like cell line (J774)," *Journal of Biological Chemistry*, vol. 268, no. 20, pp. 14896–14905, 1993.

[86] B. W. Ozanne, H. J. Spence, L. C. McGarry, and R. F. Hennigan, "Transcription factors control invasion: AP-1 the first among equals," *Oncogene*, vol. 26, no. 1, pp. 1–10, 2007.

[87] P. Verde, L. Casalino, F. Talotta, M. Yaniv, and J. B. Weitzman, "Deciphering AP-1 function in tumorigenesis: fra-ternizing on target promoters," *Cell Cycle*, vol. 6, no. 21, pp. 2632–2639, 2007.

[88] Q. Yin, X. Wang, J. McBride, C. Fewell, and E. Flemington, "B-cell receptor activation induces BIC/miR-155 expression through a conserved AP-1 element," *Journal of Biological Chemistry*, vol. 283, no. 5, pp. 2654–2662, 2008.

[89] T. Ruggiero, M. Trabucchi, F. de Santa et al., "LPS induces KH-type splicing regulatory protein-dependent processing of microRNA-155 precursors in macrophages," *Federation of American Societies for Experimental Biology Journal*, vol. 23, no. 9, pp. 2898–2908, 2009.

[90] S. L. Lin, A. Chiang, D. Chang, and S. Y. Ying, "Loss of mir-146a function in hormone-refractory prostate cancer," *RNA*, vol. 14, no. 3, pp. 417–424, 2008.

[91] A. Tenesa and M. G. Dunlop, "New insights into the aetiology of colorectal cancer from genome-wide association studies," *Nature Reviews Genetics*, vol. 10, no. 6, pp. 353–358, 2009.

[92] A. Jemal, R. Siegel, E. Ward et al., "Cancer statistics, 2006," *A Cancer Journal for Clinicians*, vol. 56, no. 2, pp. 106–130, 2006.

[93] R. J. Akhurst, "TGFβ signaling in health and disease," *Nature Genetics*, vol. 36, no. 8, pp. 790–792, 2004.

[94] R. L. Elliott and G. C. Blobe, "Role of transforming growth factor beta in human cancer," *Journal of Clinical Oncology*, vol. 23, no. 9, pp. 2078–2093, 2005.

[95] M. Langenskiöld, L. Holmdahl, P. Falk, E. V. A. Angenete, and M. L. Ivarsson, "Increased TGF-beta1 protein expression in patients with advanced colorectal cancer," *Journal of Surgical Oncology*, vol. 97, no. 5, pp. 409–415, 2008.

[96] A. Leibovitz, J. C. Stinson, and W. B. McCombs III, "Classification of human colorectal adenocarcinoma cell lines," *Cancer Research*, vol. 36, no. 12, pp. 4562–4569, 1976.

[97] J. J. Michaille, E. Tili, G. A. Calin, J. Garin, M. Louwagie, and C. M. Croce, "Cloning and characterization of cDNAs expressed during chick development and encoding different isoforms of a putative zinc finger transcriptional regulator," *Biochimie*, vol. 87, no. 11, pp. 939–949, 2005.

[98] E. Tili, J. J. Michaille, C. G. Liu et al., "GAM/ZFp/ZNF512B is central to a gene sensor circuitry involving cell-cycle regulators, TGFβ effectors, Drosha and microRNAs with opposite oncogenic potentials," *Nucleic Acids Research*, vol. 38, no. 21, pp. 7673–7688, 2010.

[99] B. N. Davis, A. C. Hilyard, G. Lagna, and A. Hata, "SMAD proteins control DROSHA-mediated microRNA maturation," *Nature*, vol. 454, no. 7200, pp. 56–61, 2008.

[100] W. Kong, H. Yang, L. He et al., "MicroRNA-155 is regulated by the transforming growth factor β/Smad pathway and contributes to epithelial cell plasticity by targeting RhoA," *Molecular and Cellular Biology*, vol. 28, no. 22, pp. 6773–6784, 2008.

[101] S. M. Wahl, J. Wen, and N. Moutsopoulos, "TGF-β: a mobile purveyor of immune privilege," *Immunological Reviews*, vol. 213, no. 1, pp. 213–227, 2006.

[102] Y. A. Yoo, M. H. Kang, J. S. Kim, and S. C. Oh, "Sonic hedgehog signaling promotes motility and invasiveness of gastric cancer cells through TGF-β-mediated activation of the ALK5-Smad3 pathway," *Carcinogenesis*, vol. 29, no. 3, pp. 480–490, 2008.

[103] K. Matsuzaki, C. Kitano, M. Murata et al., "Smad2 and Smad3 phosphorylated at both linker and COOH-terminal regions transmit malignant TGF-β signal in later stages of human colorectal cancer," *Cancer Research*, vol. 69, no. 13, pp. 5321–5330, 2009.

[104] C. Xiao, D. P. Calado, G. Galler et al., "MiR-150 controls B cell differentiation by targeting the transcription factor c-Myb," *Cell*, vol. 131, no. 1, pp. 146–159, 2007.

SAGE-Hindawi Access to Research
Journal of Nucleic Acids
Volume 2011, Article ID 131579, 6 pages
doi:10.4061/2011/131579

Research Article

Synthetic Pre-miRNA-Based shRNA as Potent RNAi Triggers

Kazuya Terasawa,[1] Kazuharu Shimizu,[2] and Gozoh Tsujimoto[1]

[1] *Department of Pharmacogenomics, Graduate School of Pharmaceutical Sciences, Kyoto University, 46-29 Yoshida Shimoadachi-cho, Sakyo-ku, Kyoto 606-8501, Japan*
[2] *Department of Nanobio Drug Discovery, Graduate School of Pharmaceutical Sciences, Kyoto University, 46-29 Yoshida Shimoadachi-cho, Sakyo-ku, Kyoto 606-8501, Japan*

Correspondence should be addressed to Gozoh Tsujimoto, gtsuji@pharm.kyoto-u.ac.jp

Received 15 October 2010; Revised 14 April 2011; Accepted 17 May 2011

Academic Editor: Dmitry A. Stetsenko

RNA interference (RNAi) is a powerful tool for studying gene function owing to the ease with which it can selectively silence genes of interest, and it has also attracted attention because of its potential for therapeutic applications. Chemically synthesized small interfering RNAs (siRNAs) and DNA vector-based short hairpin RNAs (shRNAs) are now widely used as RNAi triggers. In contrast to expressed shRNAs, the use of synthetic shRNAs is limited. Here we designed shRNAs modeled on a precursor microRNA (pre-miRNA) and evaluated their biological activity. We demonstrated that chemically synthetic pre-miRNA-based shRNAs have more potent RNAi activity than their corresponding siRNAs and found that their antisense strands are more efficiently incorporated into the RNA-induced silencing complex. Although greater off-target effects and interferon responses were induced by shRNAs than by their corresponding siRNAs, these effects could be overcome by simply using a lower concentration or by optimizing and chemically modifying shRNAs similar to synthetic siRNAs. These are challenges for the future.

1. Introduction

RNA interference (RNAi) is an evolutionarily conserved, gene-silencing mechanism that is triggered by double-stranded RNA (dsRNA). Two types of small RNA—namely, small interfering RNA (siRNA) and microRNA (miRNA)—are central players in RNAi. Both siRNAs and miRNAs regulate gene expression by annealing to mRNA sequence elements that are fully or partially complementary [1, 2]. Since transfected synthetic siRNAs were shown to induce RNAi in mammalian cells [3], they have been widely used to decipher gene function through suppression of gene expression, and they have also attracted attention because of their potential for therapeutic applications [4, 5]. miRNAs are a phylogenetically conserved family of endogenous small RNAs that play important roles in a wide variety of biological functions, including cell differentiation, tumor genesis, apoptosis, and metabolism [1, 2, 6, 7].

miRNAs are initially generated as long primary transcripts (pri-miRNA) that are processed in the nucleus by the enzyme complexes Drosha and DiGeorge Critical Region 8 (DGCR8) to a 70–90 nt stem-loop structure called pre-miRNA. The pre-miRNA is then exported to the cytoplasm. There, the exported pre-miRNA or exogenous dsRNA is cleaved by the enzyme Dicer into a ~22-nucleotide (nt) duplex known as miRNA or siRNA, respectively. The duplex is then incorporated into the RNA-induced silencing complex (RISC). After removing one strand called the passenger strand, the remaining strand, called the guide strand, in the RISC guides the silencing complex to target mRNAs. Thus, downstream of their initial processing, siRNAs and miRNAs share the same cellular machinery [1, 2].

Understanding the mechanism of the RNAi pathway has led to the development of alternative approaches of RNAi. Several groups have developed artificial miRNAs in the form of short-hairpin structures, called shRNAs, instead of siRNAs [8, 9]. DNA vector-based shRNAs are widely used. By contrast, the use of synthetic shRNAs is very limited [10–12], although synthetic shRNAs can easily incorporate chemical modifications to improve their stability and biological activity, similar to synthetic siRNAs. Probably, this limited use is because the chemical synthesis of long RNA oligonucleotides is generally difficult, consequently leading to high cost and low yield. Unlike DNA oligonucleotides, synthesis of RNA

is less efficient owing to problems caused by the presence of the 2′-hydroxyl group of ribose, which requires selective protection during oligonucleotide assembly. Recently, a new protecting group, 2-cyanoethoxymetyhyl (CEM), has been developed [13]. By improving capping and coupling conditions, the authors succeeded in synthesizing a 110 mer pre-miRNA candidate and confirmed its biological activity [14]. Thus, synthetic long oligonucleotides are becoming a reality at a reasonable cost. In this study, we have designed shRNAs modeled on pri-miRNA and have evaluated their biological activity.

2. Materials and Methods

2.1. Synthetic RNA Oligonucleotides. The siRNA sequence targeting LMNA (GenBank accession number: NM_170707) was from position 600–620 relative to the start codon. The shRNAs in Figure 1(a) were kindly provided by Nippon Shinykau Co., Ltd. (Kyoto, Japan). These RNA oligonucleotides were synthesized as previously described [14]. To anneal shRNAs, shRNAs suspended in water were incubated for 5 min at 95°C in a heat block and then left until the block reached 25°C according to the manufacturer's protocol. Control (siGENOME Non-Targeting siRNA no. 2, denoted as "ctrl") and custom designed siRNAs for LMNA were purchased from Dharmacon Research (Lafayette, CO, USA). Negative control pre-miRNA (denoted "P.N.") was from Ambion (Austin, TX, USA).

2.2. Plasmid Construction. Human Ago2 (hAgo2) cDNA was amplified by PCR using the primers 5′-GGATCCATGTACTCGGGAGCCGGC-3′ and 5′-GCGGCCGCTCAAGCAAAGTACATGGTG-3′ after reverse transcription from total RNA isolated from HeLa cells and cloned into the pCR-blunt vector (Invitrogen, Carlsbad, CA, USA). A *Kpn*I and *Not*I fragment containing a FLAG-tag coding region in the pcDNA3FLAG vector (kindly gift from E. Nishida, Kyoto University, Japan) was ligated with the *Kpn*I and *Not*I-digested expression vector pcDNA5FRT (Invitrogen), yielding the pcDNA5FLAG vector. A *BamH*I and *Not*I fragment of hAgo2 cDNA in the pCR-Blunt vector was ligated with *BamH*I and *Not*I-digested pcDNA5FLAG vector, yielding the pcDNA5FLAG–hAgo2 plasmid. An *Eco*RI site was introduced into the *Xba*I site of the luciferase reporter vector pGL4.13 (Promega, Madison, WI, USA) by ligation with the oligonucleotides 5′-CTAGACTGAATTC-3′ and 5′-CTAGGAATTCAGT-3′, yielding the pGL4.13EcoRI vector. The oligonucleotides 5′-CTAGAGAAGGAGGAACTGGACTTCCAG-3′ and 5′-AATTCTGGAAGTCCAGTTCCTCCTTCT-3′ were annealed to form a dsDNA fragment, which was ligated with *Xba*I and *Eco*RI-digested pGL4.13EcoRI to produce the pGL4-LMNA plasmid. The identity of all constructs was confirmed by DNA sequencing.

2.3. Cell Culture and Transfection. A HEK293 line stably expressing FLAG-hAgo2 was established by using the Flp-In Expression System (Invitrogen) with the pcDNA5FLAG-hAgo2 plasmid and Flp-In-293 Cell Line (Invitrogen) according to the manufacturer's protocol. HeLa cells and HEK293-derivative cells were maintained in DMEM with 10% fetal bovine serum and antibiotics (100 U/mL penicillin and 100 μg/mL streptomycin for both cell types, plus 100 μg/mL Hygromycin B for HEK293-derivative cells). Synthetic small RNAs were transfected by using X-tremeGENE siRNA Transfection Reagent (Roche, Mannheim Germany). Reporter plasmids were transfected by using FuGENE 6 Transfection Reagent (Roche). HeLa cells were plated in 24-well plates (3×10^4 cells/well) and in 12-well plates (6×10^4 cells/well), and HEK293-derivative cells were plated in 10 cm dishes (2×10^6 cells/dish) 24 h before transfection.

2.4. RNA Extraction and qRT-PCR Analysis. Total RNA was isolated from cultured cells with an RNeasy mini kit (QIAGEN, Hilden, Germany). To isolate RNA from immunoprecipitated hAgo2 protein, cells were lysed in 20 mM HEPES (pH 7.4), 150 mM NaCl, 1 mM $MgCl_2$, 1 mM EGTA, 1 mM DTT, 1% EMPIGEN BB detergent (Sigma-Aldrich, St. Louis, MO, USA), 40 U/mL R Ribonuclease Inhibitor (Takara Bio Inc., Shiga, Japan) and 1% Protease Inhibitor Cocktail (Nacalai Tesque, Kyoto, Japan). FLAG-hAgo2 was immunoprecipitated with anti-FLAG M2 agarose beads (Sigma-Aldrich), and the beads were washed three times with lysis buffer and then directly suspended in the RNA extraction reagent ISOGEN (Nippon Gene Co., Ltd., Tokyo, Japan) to elute immunoprecipitated samples containing RNAs. RNA eluted in Isogen was isolated according to the manufacturer's protocol.

To analyze mRNA expression, a QuantiTect Reverse Transcription Kit (QIAGEN) and a Power SYBR Green PCR Master Mix (Applied Biosystems, Carlsbad, CA, USA) were used. To quantify the amount of antisense and sense strand in the RISC, a miScript PCR System (QIAGEN) was used. The primers for PCR analysis were 5′-GGAAGTCCAGTTCCTCCTTC-3′ for the LMNA antisense strand, 5′-GAAGGAGGAACTGGACTTCCA-3′ for the LMNA sense strand, 5′-GAAGGAGGCTGGACTTCCA-3′ for the LMNA sense strand with a deletion, and 5′-TAGCTTATCAGACTGATGTTG-3′ for miR-21; 5′-AGTCCATTCAGACATTGGGAG-3′ and 5′-GTTGTAGATGAAGGTGAGCAG-3′ for IRF9; and 5′-CAACCATGAGTACAAATGGTG-3′ and 5′-CTAGTAGGTTGTGTATTCCCA-3′ for IFIT1.

2.5. Immunoblotting and Reporter Analysis. For the immunoblotting assay, HeLa cells plated on 12-well plates were transfected with 5 pmol of RNA oligonucleotide. After 48 h, the cells were harvested by scraping them from culture dishes into hot 1× SDS sample buffer, and the lysates were separated by SDS-PAGE and analyzed by immunoblotting. The immunoblots were visualized and quantified by using an LAS-3000 imaging system (Fujifilm, Tokyo, Japan) and normalized to the levels of β-actin. Anti-Lamin A/C rabbit IgG (no. 2032) was purchased from Cell Signaling (Danvers, MA, USA). Anti-Lamin A rabbit IgG (L1293) and β-actin monoclonal mouse IgG (clone A-15) were from Sigma-Aldrich.

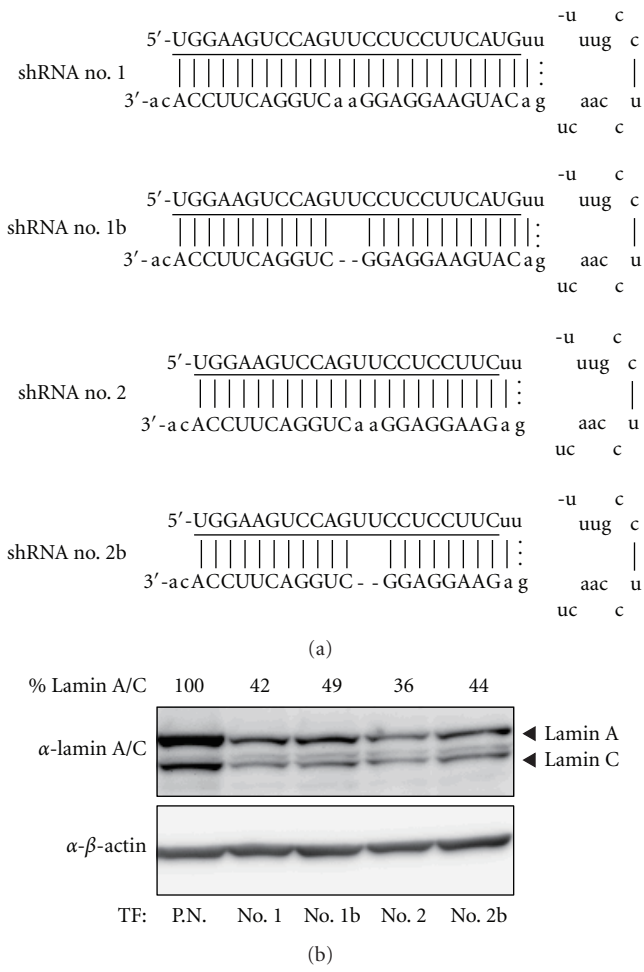


Figure 1: Design of pre-miRNA-based shRNAs and their RNAi activity. (a) The set of four synthetic shRNAs containing a 21 or 24 bp stem, either with or without a 2 bp deletion of the 3′ arm, is represented. The underlined sequences are complementary to the LMNA mRNA sequence. The lower cases are derived from the human pre-miR-155 sequence. (b) The RNAi activity of these shRNAs targeting LMNA was analyzed by immunoblotting. TF is the abbreviation for transfection.

For the reporter assay, HeLa cells plated in 24-well plates were transfected first with RNA oligonucleotide (0.25, 1.25, 2.5 pmol) 24 h before harvesting and then with reporter plasmid (100 ng/well reporter luciferase plasmid, pGL4-LMNA, and 25 ng/well renilla luciferase vector, pGL4.73; Promega) 30 min later. The luciferase activity was measured by using a dual-luciferase reporter assay system (Promega) with a Lumat LB9507 luminometer (Berthold Technologies, Bad Wildbad, Germany). As an internal control, renilla luciferase activity was used. The data reported represent the means and standard deviations of three independent experiments.

2.6. Microarray Analysis. HeLa cells plated on 12-well plates were transfected with 5 pmol of RNA oligonucleotide, and RNA was extracted 48 h after transfection. 100 ng of total RNA was amplified and labeled by using an Ambion WT Expression Kit (Ambion) according to the supplier's protocol. HG-U133 Plus 2.0 arrays (Affymetrix, Santa Clara, CA, USA) were hybridized with 11 μg of labeled cRNA, washed, stained, and scanned according to the protocol described in the Affymetrix GeneChip Expression Analysis Manual. Affymetrix data were extracted, normalized, and summarized with the robust multiaverage (RMA) method Expression Console.

3. Results

3.1. Design of Pre-miRNA-Based shRNAs. To assess the activity of synthetic shRNAs, we selected Lamin A/C (LMNA) as an RNAi target because this gene has been widely used as a positive control in siRNA experiments. Previous reports have developed miR-155-based vectors for RNAi [15, 16]; thus, we designed four shRNAs modeled on the pre-miRNA of human miR-155 (Figure 1(a)). All of the designed shRNAs had the same target site of the LMNA coding region but slightly different structures. The 17 nt loop structure and the 2 nt 3′ overhang were common parameters and are present in human pre-miR-155. It is known that Dicer efficiently cleaves dsRNA with a 2 nt 3′ overhang [11, 17] and that most pre-miRNAs have a 2 nt 3′ overhang generated by Drosha cleavage [18]. Two shRNAs (shRNA no. 1 and shRNA no. 1b) had a 24 bp stem length with or without an internal bulge, the other two (shRNA no. 2 and shRNA no. 2b) had a 21 bp stem length, generated by simply shortening the corresponding 24 bp stem shRNA at the loop side. Unlike in a previous report [11], the antisense strands were positioned 5′ to the loop. This position ensured that the antisense strand had a fixed 5′ end, irrespective of the position of cleavage by Dicer. The internal bulges in the stem region were introduced by a 2 nt deletion in the 3′-arm sequence mimic the structures of pre-miRNAs. These shRNAs were tested for their ability to silence the endogenous LMNA gene. HeLa cells were transfected with each shRNA, and the expression of LMNA protein was analyzed by immunoblotting. All four shRNAs knocked down expression of LMNA with similar efficiency (Figure 1(b)).

3.2. Potent RNAi Activity of Synthetic shRNA. Next, we compared the activity of siRNAs and shRNAs, by synthesizing siRNAs with same target sequence as the shRNAs. Immunoblotting analysis showed that shRNA no. 2 and shRNA no. 2b had RNAi activity comparable to that of the corresponding siRNA (Figure 2(a)). When shRNA no. 1 and shRNA no. 1b were used, essentially the same results were obtained (data not shown).

To analyze the activity more quantitatively, a more sensitive assay was conducted. We generated a luciferase construct that contained a perfectly complementary target site in its 3′-UTR and evaluated the reporter activity in cells cotransfected with various concentrations of synthetic small RNAs. All shRNAs, at all concentrations, were a 2- to 3-fold more potent RNAi trigger than the corresponding siRNA (Figure 2(b)). Importantly, even at the lowest concentration (0.4 nM), all shRNAs had silencing activity comparable to

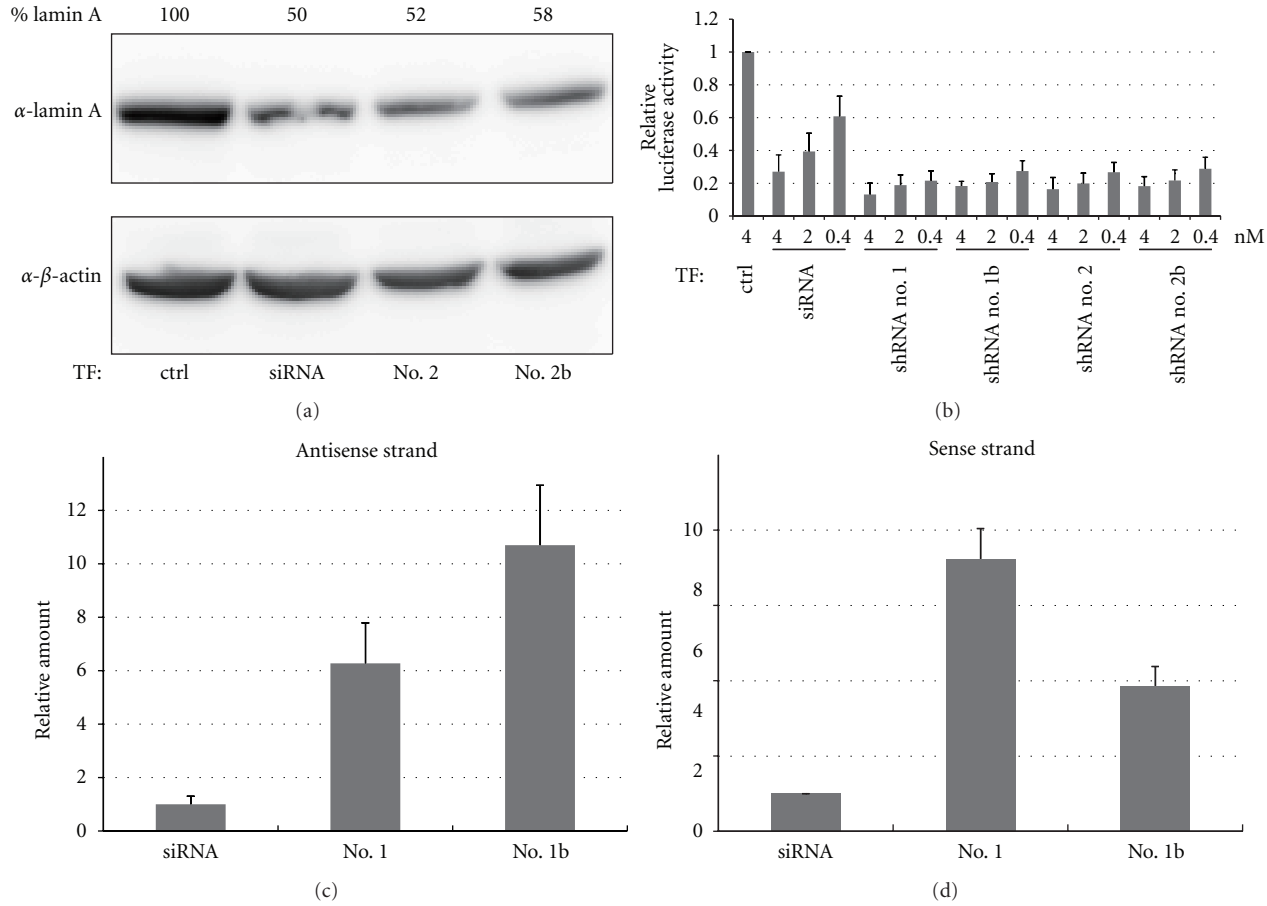


Figure 2: Comparison of the RNAi activity between synthetic shRNAs and their corresponding siRNAs. (a) The RNAi activity of siRNAs and shRNAs was analyzed by immunoblotting. (b) The RNAi activity was monitored by a reporter assay system. The luciferase activity ratio for the control siRNA (ctrl) was set as 1. (c), (d) The incorporation of transfected small RNA into the RISC was analyzed by qRT-PCR. Data were normalized against the amount of miR-21 in the RISC. The amount of siRNA incorporated was set as 1. Each graph point in (b), (c), and (d) represents the average (with s.d.) of three independent experiments.

that observed at their 10-fold higher concentration (4 nM). These results are coincident with previous reports [11, 19] and indicate that synthetic shRNAs are more potent than their corresponding siRNAs. In this assay, we could not detect significant differences among the four shRNAs tested.

3.3. Efficiently Incorporation of Synthetic shRNAs into the RISC. As previously discussed [11], we reasoned that the more effective RNAi achieved by shRNAs as compared with siRNAs might reflect a difference in the efficiency of their incorporation into the RISC. We therefore examined the amount of antisense strands in the RISC generated from transfected small RNAs. Two shRNAs (no. 1 and no. 1b) were transfected into HEK293 cells stably expressing FLAG-tagged hAgo2 protein, which is a core component of the RISC [20]. Processed RNAs loaded into the RISC were isolated by coimmunoprecipitation with FLAG-tagged hAgo2 protein. As expected, real-time quantitative RT-PCR analysis revealed that the antisense strands derived from shRNAs were more efficiently incorporated into the RISC as compared with those derived from the corresponding siRNA (Figure 2(c)). Presumably, this efficient incorporation into the RISC underlies the highly potent silencing activity of shRNAs.

Interestingly, introducing an internal bulge in the stem increased the amount of antisense strands (compare no. 1b with no. 1 in Figure 2(c)). This might indicate that the equilibrium between the sense and antisense strands shifted towards antisense loading. We tested this equilibrium shift by quantifying the amount of antisense strands in the RISC. Indeed, introducing an internal bulge changed the equilibrium toward antisense loading (Figure 2(d)).

3.4. Off-Target Effects and Interferon Responses Induced by Synthetic shRNAs in HeLa Cells. It is known that siRNAs can cause off-target effects [21]. Using microarray gene expression profiling, we compared the off-target effects of our synthetic small RNAs. From gene expression profiling, we identified genes that showed a more than 2-fold decrease in expression as compared with the control siRNA transfected sample. A Venn diagram of these downregulated genes is shown in Figure 3. As compared with siRNA, transfection

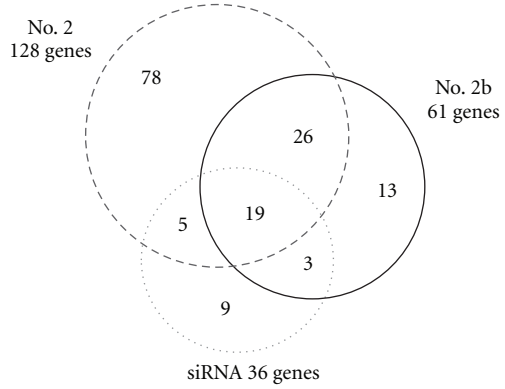


Figure 3: Off-target effects induced by synthetic shRNAs. HeLa cells were transfected with the indicated RNAs. After 48 h, the expression of about 20,000 protein-coding genes was analyzed by microarray analysis. The overlap of genes that showed a more than 2-fold decrease in expression as compared with the control siRNA-transfected sample is shown. The number in the overlapping region of the Venn diagram represents shared genes. Note that the LMNA gene is included in the commonly downregulated genes.

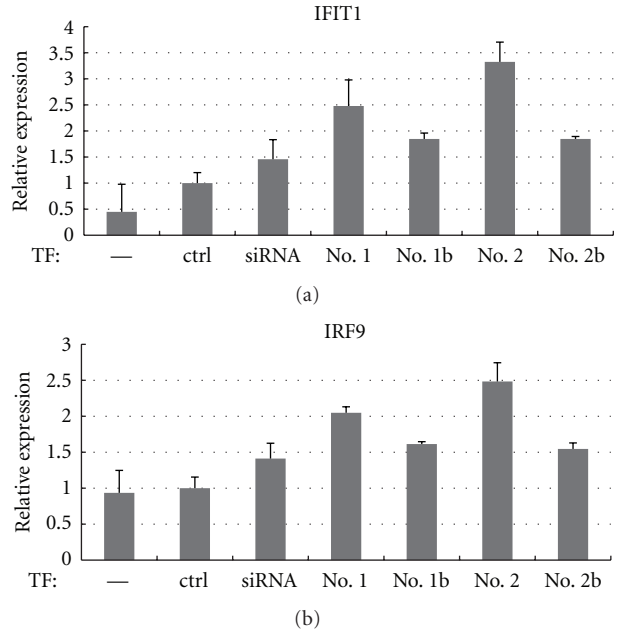


Figure 4: Interferon responses induced by synthetic shRNAs. HeLa cells were transfected with the indicated RNAs. After 48 h, the expression of two interferon-regulated genes IFIT1 and IRF9 was analyzed by qRT-PCR.

of shRNAs caused a broader downregulation of nontargeted transcripts. This might be simply because of the efficient incorporation of synthetic shRNAs into the RISC and the consequent higher potent RNAi activity (Figure 2(c)). We noted that introducing a bulge decreased the number of off-target genes (compared no. 2b with no. 2). This phenomenon might reflect preferential incorporation of the antisense strand into the RISC, as occurred when shRNAs with a bulge were used (Figure 2(d)).

Next, we examined the expression of interferon-related genes among the microarray data and found that IFIT1 [22] and IRF9 [23] were induced more than 2-fold by transfection of shRNA no. 2. We further analyzed the expression of these two genes by qRT-PCR analysis (Figure 4). Transfection of the siRNA upregulated IRF9 and IFIT1 expression by approximately 1.5-fold; however, transfection of shRNAs without a bulge caused slightly higher upregulation of these genes. Again, shRNAs with a bulge improved interferon responses. IFIT1 is known to be the mRNA that is most strongly induced in response to dsRNAs [22]. Introducing a bulge into the stem could circumvent this immune activation.

4. Discussion

In this study, we demonstrated that synthetic pre-miRNA-based shRNAs have more potent RNAi activity than their corresponding siRNA and found that their antisense strands are more efficiently incorporated into the RISC. A previous study also showed that synthetic shRNAs with a 27 nt stem and a 4 nt artificial loop have a significantly higher gene-silencing activity than conventional 21 nt siRNAs [11]. The authors demonstrated that these small RNAs are subjected to Dicer processing both *in vitro* and *in vivo* and speculated that siRNA duplexes generated by Dicer are efficiently loaded into the RISC. In a study using DNA vector-based shRNAs, two types of shRNA construct modeled on pre-miRNA and pri-miRNA transcripts were examined [24]. The results indicated that RNAi triggers that enter the RNAi pathway by a more natural route yield more effective silencing [25]. Here we demonstrated the efficient incorporation of shRNAs into the RISC, and our results are highly consistent with those of previous studies. Furthermore, introducing a bulge enhanced incorporation into the RISC and shifted the equilibrium toward antisense loading (Figures 2(c) and 2(d)). However, we did not observe a significant difference in RNAi activity among the synthetic shRNAs in our assay (Figure 2(b)), probably because they were all close to saturation.

There are contradictory reports about the effects of stem length on RNAi activity. Some reports have shown that shRNAs with a 19 nt stem and a 9 or 10 nt loop have higher RNAi activity than shRNAs with longer stem [12, 26, 27]. This type of shRNA is widely used as a vector-based shRNA [28]. Interestingly, shRNAs with a 19 nt stem are not Dicer substrates; however, these shRNAs are thought to be incorporated into the RISC after bypassing Dicer processing [11]. Although we reasoned that the development of shRNAs based on naturally found structures is a promising approach, we should also examine and evaluate this type of shRNA.

Unlike previously used shRNAs [11], our designed shRNAs have a shorter stem length (21 or 24 nt) and the loop derived from a pre-miRNA and are therefore expected to be less toxic. However, we observed that interferon responses were induced more strongly by shRNAs than by their corresponding siRNA. Interestingly, introducing a bulge also

reduced interferon responses. These observations, albeit of a few examples, indicate that introducing a bulge has potential for improving the performance of shRNAs. Further examination of the position and structure of bulges is required. Chemical modifications of siRNAs have been extensively studied to increase stability, promote efficacy, minimize off-target effects, and reduce innate immune responses [5, 21]. Such modifications can be easily incorporated in synthetic shRNAs to improve siRNA stability and biological activity, similar to synthetic siRNAs. These are issues to be addressed in the future.

Acknowledgments

The authors are grateful to their laboratory members for technical assistance and fruitful discussion and comments. This work was supported in part by grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan (no. 19109001 to G. Tsujimoto and K. Terasawa, and no. 20390342 to G. Tsujimoto and K. Shimizu), and the New Energy and Industrial Technology Development Organization (to G.T and K.T).

References

[1] M. Ghildiyal and P. D. Zamore, "Small silencing RNAs: an expanding universe," *Nature Reviews Genetics*, vol. 10, no. 2, pp. 94–108, 2009.

[2] R. W. Carthew and E. J. Sontheimer, "Origins and Mechanisms of miRNAs and siRNAs," *Cell*, vol. 136, no. 4, pp. 642–655, 2009.

[3] S. M. Elbashir, J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, and T. Tuschl, "Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells," *Nature*, vol. 411, no. 6836, pp. 494–498, 2001.

[4] P. D. Zamore, "RNA interference: big applause for silencing in Stockholm," *Cell*, vol. 127, no. 6, pp. 1083–1086, 2006.

[5] D. Castanotto and J. J. Rossi, "The promises and pitfalls of RNA-interference-based therapeutics," *Nature*, vol. 457, no. 7228, pp. 426–433, 2009.

[6] V. Ambros, "The functions of animal microRNAs," *Nature*, vol. 431, no. 7006, pp. 350–355, 2004.

[7] S. Tsuchiya, Y. Okuno, and G. Tsujimoto, "MicroRNA: biogenetic and functional mechanisms and involvements in cell differentiation and cancer," *Journal of Pharmacological Sciences*, vol. 101, no. 4, pp. 267–270, 2006.

[8] P. J. Paddison, A. A. Caudy, E. Bernstein, G. J. Hannon, and D. S. Conklin, "Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells," *Genes and Development*, vol. 16, no. 8, pp. 948–958, 2002.

[9] Y. Zeng, E. J. Wagner, and B. R. Cullen, "Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells," *Molecular Cell*, vol. 9, no. 6, pp. 1327–1333, 2002.

[10] M. T. McManus, C. P. Petersen, B. B. Haines, J. Chen, and P. A. Sharp, "Gene silencing using micro-RNA designed hairpins," *RNA*, vol. 8, no. 6, pp. 842–850, 2002.

[11] D. Siolas, C. Lerner, J. Burchard et al., "Synthetic shRNAs as potent RNAi triggers," *Nature Biotechnology*, vol. 23, no. 2, pp. 227–231, 2005.

[12] A. V. Vlassov, B. Korba, K. Farrar et al., "shRNAs targeting hepatitis C: effects of sequence and structural features, and comparison with siRNA," *Oligonucleotides*, vol. 17, no. 2, pp. 223–236, 2007.

[13] T. Ohgi, Y. Masutomi, K. Ishiyama, H. Kitagawa, Y. Shiba, and J. Yano, "A new RNA synthetic method with a 2′-O-(2-cyanoethoxymethyl) protecting group," *Organic Letters*, vol. 7, no. 16, pp. 3477–3480, 2005.

[14] Y. Shiba, H. Masuda, N. Watanabe et al., "Chemical synthesis of a very long oligoribonucleotide with 2-cyanoethoxymethyl (CEM) as the 2′-O-protecting group: structural identification and biological activity of a synthetic 110mer precursor-microRNA candidate," *Nucleic Acids Research*, vol. 35, no. 10, pp. 3287–3296, 2007.

[15] K. H. Chung, C. C. Hart, S. Al-Bassam et al., "Polycistronic RNA polymerase II expression vectors for RNA interference based on BIC/miR-155," *Nucleic Acids Research*, vol. 34, no. 7, article e53, 2006.

[16] J. Wu, A. N. Bonsra, and G. Du, "pSM155 and pSM30 vectors for miRNA and shRNA expression," *Methods in Molecular Biology*, vol. 487, pp. 205–219, 2009.

[17] H. Zhang, F. A. Kolb, L. Jaskiewicz, E. Westhof, and W. Filipowicz, "Single processing center models for human Dicer and bacterial RNase III," *Cell*, vol. 118, no. 1, pp. 57–68, 2004.

[18] Y. Lee, C. Ahn, J. Han et al., "The nuclear RNase III Drosha initiates microRNA processing," *Nature*, vol. 425, no. 6956, pp. 415–419, 2003.

[19] D. H. Kim, M. A. Behlke, S. D. Rose, M. S. Chang, S. Choi, and J. J. Rossi, "Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy," *Nature Biotechnology*, vol. 23, no. 2, pp. 222–226, 2005.

[20] L. Peters and G. Meister, "Argonaute proteins: mediators of RNA silencing," *Molecular Cell*, vol. 26, no. 5, pp. 611–623, 2007.

[21] A. L. Jackson and P. S. Linsley, "Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application," *Nature Reviews Drug Discovery*, vol. 9, no. 1, pp. 57–67, 2010.

[22] G. Geiss, G. Jin, J. Guo, R. Bumgarner, M. G. Katze, and G. C. Sen, "A comprehensive view of regulation of gene expression by double-stranded RNA-mediated cell signaling," *Journal of Biological Chemistry*, vol. 276, no. 32, pp. 30178–30182, 2001.

[23] S. A. Veals, C. Schindler, D. Leonard et al., "Subunit of an alpha-interferon-responsive transcription factor is related to interferon regulatory factor and Myb families of DNA-binding proteins," *Molecular and Cellular Biology*, vol. 12, no. 8, pp. 3315–3324, 1992.

[24] J. M. Silva, M. Z. Li, K. Chang et al., "Second-generation shRNA libraries covering the mouse and human genomes," *Nature Genetics*, vol. 37, no. 11, pp. 1281–1288, 2005.

[25] B. R. Cullen, "RNAi the natural way," *Nature Genetics*, vol. 37, no. 11, pp. 1163–1165, 2005.

[26] L. M. Li, X. Y. Lin, A. Khvorova, S. W. Fesik, and Y. Shen, "Defining the optimal parameters for hairpin-based knockdown constructs," *RNA*, vol. 13, no. 10, pp. 1765–1774, 2007.

[27] Q. Ge, H. Ilves, A. Dallas et al., "Minimal-length short hairpin RNAs: the relationship of structure and RNAi activity," *RNA*, vol. 16, no. 1, pp. 106–117, 2010.

[28] T. R. Brummelkamp, R. Bernards, and R. Agami, "A system for stable expression of short interfering RNAs in mammalian cells," *Science*, vol. 296, no. 5567, pp. 550–553, 2002.

SAGE-Hindawi Access to Research
Journal of Nucleic Acids
Volume 2011, Article ID 246127, 14 pages
doi:10.4061/2011/246127

Research Article

Conformational-Dependent and Independent RNA Binding to the Fragile X Mental Retardation Protein

Xin Yan[1] and Robert B. Denman[2]

[1] *CSI/IBR Center for Developmental Neuroscience, College of Staten Island, City University of New York, Staten Island, NY 10314, USA*
[2] *Biochemical Molecular Neurobiology Laboratory, Department of Molecular Biology, New York State Institute for Basic Research in Developmental Disabilities, 1050 Forest Hill Road, Staten Island, NY 10314, USA*

Correspondence should be addressed to Robert B. Denman, rbdenman@yahoo.com

Received 8 October 2010; Accepted 16 March 2011

Academic Editor: Dmitry A. Stetsenko

The interaction between the fragile X mental retardation protein (FMRP) and BC1 RNA has been the subject of controversy. We probed the parameters of RNA binding to FMRP in several ways. Nondenaturing agarose gel analysis showed that BC1 RNA transcripts produced by in vitro transcription contain a population of conformers, which can be modulated by preannealing. Accordingly, FMRP differentially binds to the annealed and unannealed conformer populations. Using partial RNase digestion, we demonstrate that annealed BC1 RNA contains a unique conformer that FMRP likely binds. We further demonstrate that this interaction is 100-fold weaker than that the binding of eEF-1A mRNA and FMRP, and that preannealing is not a general requirement for FMRP's interaction with RNA. In addition, binding does not require the N-terminal 204 amino acids of FMRP, methylated arginine residues and can be recapitulated by both fragile X paralogs. Altogether, our data continue to support a model in which BC1 RNA functions independently of FMRP.

1. Introduction

Fragile X syndrome is the most common inherited cause of mental impairment accounting for *ca.* 40% of X-linked mental retardation cases. It is also the most common known cause of autism (reviewed in [1–6]). Other characteristics of the fragile X syndrome include hyperactivity [7], increased susceptibility to seizures [8], increased testicular volume [9], macrocephaly, and large ears [10]. In addition, it has been found that carriers of the fragile X premutation, once thought to be free of the effects of the disease, also suffer from subtle behavioral and physical abnormalities [11–14]. This wide and varied constellation of phenotypic features results from the loss of function of a single gene, *FMR1* (summarized in: http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=gene&part-fragilex).

The *FMR1* gene encodes the RNA-binding protein FMRP [15], a negative [16–18] and positive [19, 20] translational regulator, and it has been of considerable interest to delineate the cellular RNAs that bind to FMRP [21–25] and the mechanism(s) by which FMRP binds and controls these mRNAs [26–35].

In 2003, Zalfa et al. described a bridging mechanism in which the fragile X mental retardation protein (FMRP) *via* interaction with the 5′ end of the small noncoding RNA, BC1, and bound and repressed FMRP target mRNAs [36]. This model has been subject to great deal of scrutiny owing to findings that appear to be out of step with other studies. These include differences in the prime localization of FMRP with small repressed mRNPs rather than brain polyribosomes [18, 37–39], differences in the interpretation of the interaction of FMRP with BC1 RNA as specific and significant [40] rather than nonspecific and insignificant [30, 41, 42] and detailed mechanistic differences in the nature of BC1 RNA-mediated localization and translational repression [41, 43–45]. In response to some of these criticisms, Zalfa

and Bagni reposited that their model, rather than being a *general* model of FMRP-mRNA interactions, was only one of several possible models [46].

More recent investigations of some of the concomitants of the Zalfa model [47] found that the interaction of recombinant FMRP with BC1 RNA was weak compared to that of a G-quartet-containing RNA. Furthermore, its strength varied significantly depending upon the buffer conditions used. The data suggested that FMRP may interact with a particular conformer of BC1 RNA. Here, we elaborate the conditions and requirements for a weak FMRP BC1 RNA in vitro interaction.

2. Materials and Methods

2.1. Buffers. Z-buffer is 10 mM Tris-HCl pH 7.5, 2 mM $MgCl_2$, 400 mM NaCl and 0.2% SDS [36]. The RNA-binding buffer that was used for affinity capture is 50 mM Tris-HCl, pH 7.4, 2 mM $MgCl_2$, and 150 mM KCl, 1 mM EDTA, and 1 mM DTT [26]. Structure buffer is 10 mM Tris-HCl pH 7, 100 mM KCl, and 10 mM $MgCl_2$.

2.2. Preparation of BC1 RNA Transcripts. *Dra I* linearized pBCX607 containing the entire BC1 sequence [48], *AvaII* linearized pBCX607 containing the first 65 b of BC1, *Sac I* linearized pMK-1 containing the last 60 b of BC1 RNA, a PCR fragment encoding a T7 RNA polymerase promoter and the first 75 b of BC1 RNA, *Hind III* linearized pTAR encoding an 85 b transcript that folds into a 57 b TAR element and a 28 b leader sequence and linearized pTri-XEF1 encoding eEF-1A mRNA (Ambion) were used to produce biotinylated RNAs via *in vitro* transcription (Ambion). Plasmids pBCX607 and pMK-1 were provided by Dr. Henri Tiedge (SUNY Brooklyn). Alcohol-precipitated RNAs were dissolved in 50 μL DEPC-treated H_2O and quantified spectrophotometrically. RNA integrity was examined by agarose gel electrophoresis. For examining the effect of annealing on RNA-protein interactions, individual RNAs were heated in 1x transcription buffer (Ambion) at 65°C for 10 min and then allowed to cool slowly to room temperature for one hour prior to their use. Note. We obtained identical results using Z-buffer in the annealing reaction (not shown).

2.3. RNA-Binding Assays. ^{35}S-FMRP, ^{35}S-$FMRP_{280}$, ^{35}S-$FMRP_{204}$, ^{35}S-FXR1P, ^{35}S-FXR2P, ^{35}S-eIF4A, and ^{35}S-luciferase and were produced from plasmids pET21A-FMRP, pND-L-mHisFMRP, pET9-$FMRP_{280}$, pET9-$FMRP_{204}$, pHA-FXR1P, and pET21b-FXR2P in an RRL-coupled transcription-translation system (Promega). Plasmid pHA-FXR1P was provided by Dr. Gideon Dreyfuss (University of Pennsylvania); plasmids pET9-$FMRP_{280}$ and pET9-$FMRP_{204}$ were provided by Dr. Darryl Spinner (IBR), plasmid pET21b-FXR2P was a gift from Dr. Jennifer Darnell (Rockefeller University) and plasmid pET-His6-eIF4A was a gift from Dr. Henri Tiedge (SUNY Brooklyn). Briefly, ^{35}S-labeled-proteins were produced by combining twenty five microliters of TNT rabbit reticulocyte lysate (RRL) with 2 μL of TNT T7 RNA polymerase, 1 μL of 1 mM amino acid mix minus-methionine, 35 μCi of ^{35}S-methionine, 1 μL of RNasin and 1 μg of plasmid DNA in a total volume of 50 μL. One microliter of a 50X Complete protease inhibitor cocktail was added to prevent proteolysis. Samples were incubated at 30°C for 90 min and then assayed for protein production by autoradiography.

Affinity capture assays were performed as described previously [17]; the bound and unbound products were resolved by SDS-PAGE and subject to autoradiography.

Autoradiograms were quantified using UN-SCAN-IT Gel 6.1 (Silk Scientific, Inc.). The percent binding was calculated as

$$\%\text{Binding} = 100 \times \frac{\text{Intensity}_{\text{bound}}}{\left[\text{Intensity}_{\text{bound}} + \text{Intensity}_{\text{unbound}}\right]}. \quad (1)$$

The percent binding of the "no RNA" control in each experimental set was subtracted from that of the samples; the difference, representing authentic binding, was plotted.

2.4. RNA Structure Studies. Annealed and unannealed BC1 RNA, BC1 fragment RNAs and control RNAs (1-2 μg), were treated at room temperature for 15 min with various amounts of ribonuclease V1 (cobra venom) or ribonuclease A in structure buffer as indicated. Reaction products were resolved on 1-2% TAE agarose gels containing 0.1 μg/mL ethidium bromide along with appropriate size markers. Gels were imaged using a Scion CFW-1308 M mega pixel camera and captured in inverted mode using FOTO/Analyst PC Image software version 9.04 (FOTODYNE). The resulting image files were digitized and analyzed using UN-SCAN-IT Gel 6.1.

2.5. RNA Secondary Structure Modeling. RNA lowest energy secondary structures were determined using the Zuker algorithm, M-fold (http://mfold.rna.albany.edu).

3. Results and Discussion

3.1. Affinity Capture of FMRP with Biotinylated RNA. Several methods have been used to assess the direct physical interaction of FMRP with RNA *in vitro*. These include pull-down assays with homoribopolymers [15, 21, 49, 50], affinity capture using biotinylated RNA [21, 41, 51–53], affinity capture using immobilized protein [16], UV crosslinking [21, 49], filter-binding assays [22, 30, 41], electrophoretic mobility shift assays (EMSA) [26, 36, 40, 42, 47, 54], and agarose electrophoretic mobility shift assays (AGESA) [17, 41]. Each of these methods has its unique experimental advantages [55]. Acknowledging that binding between a nucleic acid and RNA-binding protein (RBP) can be affected by differences in posttranslational modification [56] and/or differences between different protein variants [57, 58] our working hypothesis is that given a particular RBP, a particular RNA and a defined buffer each of these methods should converge to produce a common answer. While extensively studied, FMRP's interactions with RNA have not always been examined with this hypothesis in mind Table 1, and it has been suggested that differences in experimental conditions

Table 1: Binding conditions used to measure the interaction between FMRP and RNA.

Publication	Assay	Conditions
Ashley et al. (1993) [51]	Pull-down	16 mM HEPES-KOH pH 7.9, 120 mM KCI, 0.04% Nonidet P-40, 1 mg/mL BSA, 0.16 mM dithioerythritol, 0.4 mM phenylmethylsulfonyl fluoride
Brown et al. (1998) [49]	Pull-down	10 mM Tris-HCl pH 7.5, 2.5 mM $MgCl_2$, 100 mM NaCl, 2.5% Trition X100, 1 mg/mL heparin
Price et al. (1996) [52]	Pull-down	20 mM Hepes, pH 7.9, 2 mM $MgCl_2$, 10 mM $ZnCl_2$, 70 mM NH_4Cl, 0.02% Nonidet P-40, 5 mg/mL yeast tRNA
Sung et al. (2000) [21]	Pull-down Filter-Binding	20 mM Hepes, pH 7.9, 2 mM $MgCl_2$, 10 mM $ZnCl_2$, 70 mM NH_4Cl, 0.02% Nonidet P-40, 5 mg/mL yeast tRNA
Denman and Sung (2002) [57]	Pull-down	20 mM Hepes, pH 7.9, 2 mM $MgCl_2$, 10 mM $ZnCl_2$, 70 mM NH_4Cl, 0.02% Nonidet P-40, 5 mg/mL yeast tRNA
Schaeffer et al. (2001) [26]	EMSA	50 mM Tris-HCl pH 7.4, 1 mM $MgCl_2$, 1 mM EDTA, 150 mM KCl, 1 mM DTT, 0.25 mg/mL of *E.coli* tRNA, 0.01% BSA, 8 U of RNasin
Sung et al. (2003) [17]	Pull-down EMSA	50 mM Tris-HCl, pH 7.0, 2 mM $MgCl_2$, 150 mM NaCl, 1 mM DTT, 0.25 mg/mL *E.coli* tRNA, 0.25 mg/mL BSA
Bechara et al. (2006) [61]	EMSA	50 mM Tris-HCl pH 7.4, 1 mM $MgCl_2$, 1 mM EDTA, 150 mM KCl, 1 mM DTT, 0.25 mg/mL of *E.coli* tRNA, 0.01% BSA, 8 U of RNasin
Didiot et al. (2008) [54]	EMSA	50 mM Tris-HCl pH 7.4, 1 mM $MgCl_2$, 1 mM EDTA, 150 mM KCl, 1 mM DTT, 0.25 mg/mL of *E. coli* tRNA, 0.01% BSA, 8 U of RNasin
Zalfa et al. (2003) [36]	EMSA	10 mM HEPES pH 7.9, 3 mM $MgCl_2$, 10 mM DTT, 100 mM KCl, 750 mM NaCl, 5% glycerol, 7 mM β-Mercaptoethanol, 1 mg/mL Albumin, 1.3 mg/mL Heparin
Zalfa et al. (2005) [40]	EMSA	20 mM HEPES-KOH, pH 7.6, 5 mM $MgCl_2$, 300 mM KCl, 2 mM DTT, 5% glycerol, and 500 ng of total yeast tRNA or 20 μg of heparin.
Darnell et al. (2001) [22]	Filter-Binding	10 mM Tris-OAc pH 7.7, 200 mM KOAc, 5 mM $MgOAc_2$
Darnell et al. (2005) [30]	Filter-Binding	50 mM Tris-OAc at pH 7.7, 50 mM KOAc, 10 mM DTT, 5 mM $Mg(OAc)_2$, 30 μg/mL tRNA
Gabus et al. (2004) [42]	EMSA	20 mM Tris-HCl pH 7.5, 30 mM NaCl, 0.2 mM $MgCl_2$, 5 mM DTT, 10 μM $ZnCl_2$
Laggerbauer et al. (2001) [16]	Pull-down	PBS, 0.02% IGEPAL, 1% BSA
Siomi et al. (1993) [15]	Pulldown	10 mM Tris-HCl pH 7.4, 2.5 mM $MgCl_2$, 0.5% Triton X-100, 100–1000 mM NaCl
Stetler et al. (2005) [56]	Pulldown	2 M KOAc, 100 mM Tris-OAc pH 7.7 and 50 mM $MgOAc_2$, 1 μL of yeast tRNA, 1 μL of RNAsin
Menon and Mihailescu (2007) [62]	EMSA	50 mM Tris-HCl pH 7.5, 150 mM NaCl and protease inhibitors
Fahling et al. (2009) [20]	EMSA	10 mM Hepes pH 7.2, 3 mM $MgCl_2$, 5% glycerol, 1 mM DTT, 150 mM KCl, 2 U/μL RNaseOUT, 0.5 μg/μL rabbit rRNA
Zou et al. (2008) [63]	AGESA	20 mM Tris-HCl pH 7.2, 150 mM NaCl
Iacoangeli et al. (2008) [47]	EMSA	50 mM Tris-HCl pH 7.6, 150 mM KCl, 1 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT, 0.2 U/μL RNase inhibitor, 100 ng/μL total yeast tRNA, and 100 ng/μL BSA

Sets of binding conditions are grouped by the different laboratories that used them. Each set of conditions is differentiated from the next by a dotted line. In some instances the same group used multiple sets of binding conditions in multiple publications.

and protein preparations form the basis for the divergent results obtained for the FMRP BC1 RNA interaction [59, 60].

We chose to examine heretofore unstudied aspects of FMRP's interaction with BC1 RNA-using affinity capture, one of the older techniques used in FMRP RNA binding studies. We next sought a basic buffer to use. In several previous publications, we have used a buffer described by Schaeffer et al. [26], which used physiological saline [17, 41, 47], Table 1. However, this buffer contains tRNA to reduce nonspecific binding, and as two publications demonstrate that FMRP can interact with tRNA [42, 47], we first determined whether specific binding between FMRP and RNA could be observed without the addition of tRNA. As shown in Figure 1 (upper panel), in the presence of physiological salt and in the absence of RNA, ^{35}S-FMRP produced by *in vitro* translation in rabbit reticulocyte lysate (RRL) nonspecifically bound to the avidin affinity column; however, addition of increasing amounts of NaCl decreased this nonspecific interaction so that at 125 mM NaCl the amount of bound FMRP was between 5%–10% of the total. As expected, tRNA also blocked the association of full-length FMRP to the avidin column, Figure 1 (lower panel). Thus,

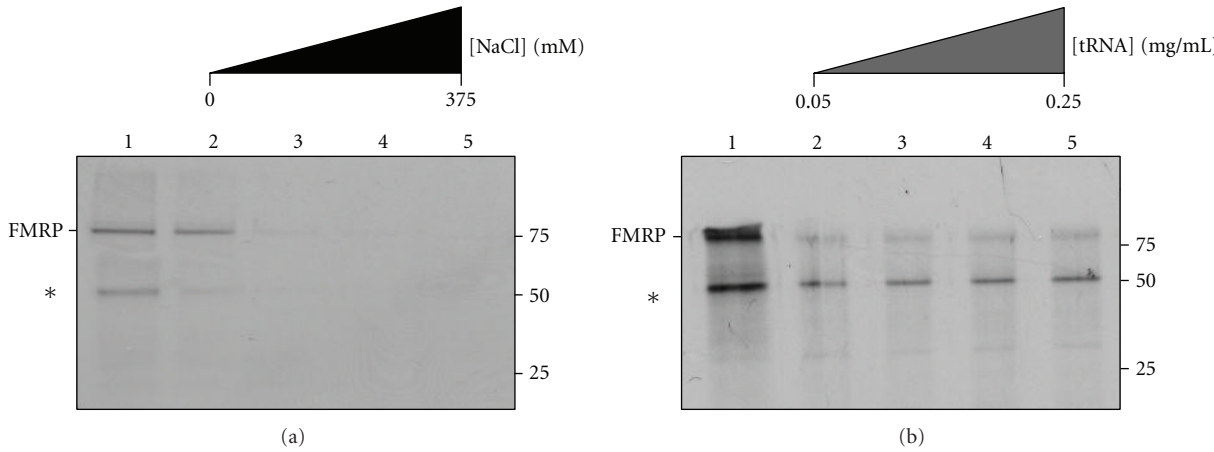


Figure 1: NaCl blocks nonspecific binding of FMRP with avidin. (a) ^{35}S-FMRP was bound to SoftLink resin in 1x Schaeffer binding buffer supplemented with 0 mM, 125 mM, 250 mM and 375 mM NaCl, lanes 2–5, respectively. Bound ^{35}S-FMRP was recovered, resolved by SDS-PAGE and subject to autoradiography. Lane 1 shows the amount of ^{35}S-FMRP input into the assay. (b) ^{35}S-FMRP was bound to SoftLink resin in 1x Schaeffer binding buffer supplemented with 0.05 mg/mL, 0.1 mg/mL, 0.15 mg/mL, and 0.25 mg/mL tRNA, lanes 2–5, respectively. Bound ^{35}S-FMRP was recovered, resolved by SDS-PAGE and subject to autoradiography. Lane 1 shows the amount of ^{35}S-FMRP input into the assay. The asterisk marks a ^{35}S-truncation product produced by transcription/translation.

Schaeffer buffer supplemented with 125 mM NaCl blocks the non-specific interaction of FMRP with the affinity matrix as effectively as Schaeffer buffer supplemented with tRNA.

3.2. Prior Annealing of BC1 RNA Enhances Its Interaction with FMRP. Having established this basic set of conditions, we applied them to study the FMRP BC1 RNA interaction. The weak binding between FMRP and BC1 RNA that occurs in protein excess [36] suggested that the protein might be surveying the population of BC1 RNA conformers and interacting with a particular one. As BC1 RNA is known to form higher order structure [45], we performed a simple experiment to perturb the conformer population that would test this hypothesis. BC1 RNA was transcribed *in vitro* and purified by salt and alcohol precipitation; then, it was either heated briefly or left untreated. Each RNA (annealed and unannealed, resp.) was then was titrated with a constant amount of ^{35}S-FMRP, and the binding was assessed by affinity capture. Under the conditions used, unannealed BC1 RNA bound extremely weakly at all concentrations examined, in concert with previous obtained results [41, 64]. On the other hand, annealed BC1 RNA exhibited stronger binding to FMRP over the range of concentrations examined, Figures 2(a) and 2(b).

To index the above results to a known standard, the interaction between eEF-1A mRNA and FMRP was also measured in parallel. Previous work had shown that this RNA binds strongly to FMRP, without prior annealing [17, 41]. As expected, the FMRP eEF-1A mRNA interaction was much stronger than the FMRP BC1 RNA interaction, Figure 2(c). In fact, with the same amount of ^{35}S-FMRP, unannealed eEF-1A mRNA evinced saturable binding at 90 nM RNA, while it took 100-fold more annealed BC1 RNA to achieve comparable binding; see Figure 2(d).

To determine whether annealing affected the interaction of FMRP with eEF-1A mRNA, binding to annealed and unannealed forms of the message was assessed using a subsaturating concentration of eEF-1A RNA. As shown in Figure 2(e), binding of eEF-1A mRNA to ^{35}S-FMRP was not markedly affected by annealing.

3.3. Annealing Alters the Structure of BC1 RNA. The above data implied that the conformer populations of annealed and unannealed BC1 RNA differ. To determine whether this was so, the annealed and unannealed BC1 RNAs used in Figure 2 were treated with RNases whose cleavage depends on known RNA structural features and the products were resolved on nondenaturing agarose gels. In the absence of treatment, both RNAs displayed a major band as well as a less prominent, slower migrating band, Figure 3(b). These data indicate that annealed BC1 RNA and unannealed BC1 RNA contained multiple conformers. Interestingly, in the presence of RNase A, which preferentially cleaves at single-stranded C and U residues, both annealed and unannealed BC1 RNA behaved nearly identically and were completely degraded by all but the lowest amount of the enzyme. However, when the RNAs were treated with a range of concentrations of RNase V1 (0.01–1 units), which cleaves base-paired nucleotides, the unannealed form was refractory to cleavage, while the annealed form was sensitive toward cleavage at the highest amount, Figure 3(c) (upper panel). Increasing the amount of RNase V1 recapitulated the sensitivity of annealed BC1 RNA, Figure 3(c) (lower panel); however, it also demonstrated that a fraction of unannealed BC1 RNA also contained RNase V1-sensitive stable duplex RNA. To determine whether the unique BC1 RNA conformer(s) produced by annealing could be stabilized by Mg^{+2} the annealing reaction was also carried out in the presence of 2 mM $MgCl_2$. The results indicated that the magnesium did not alter the distribution of BC1

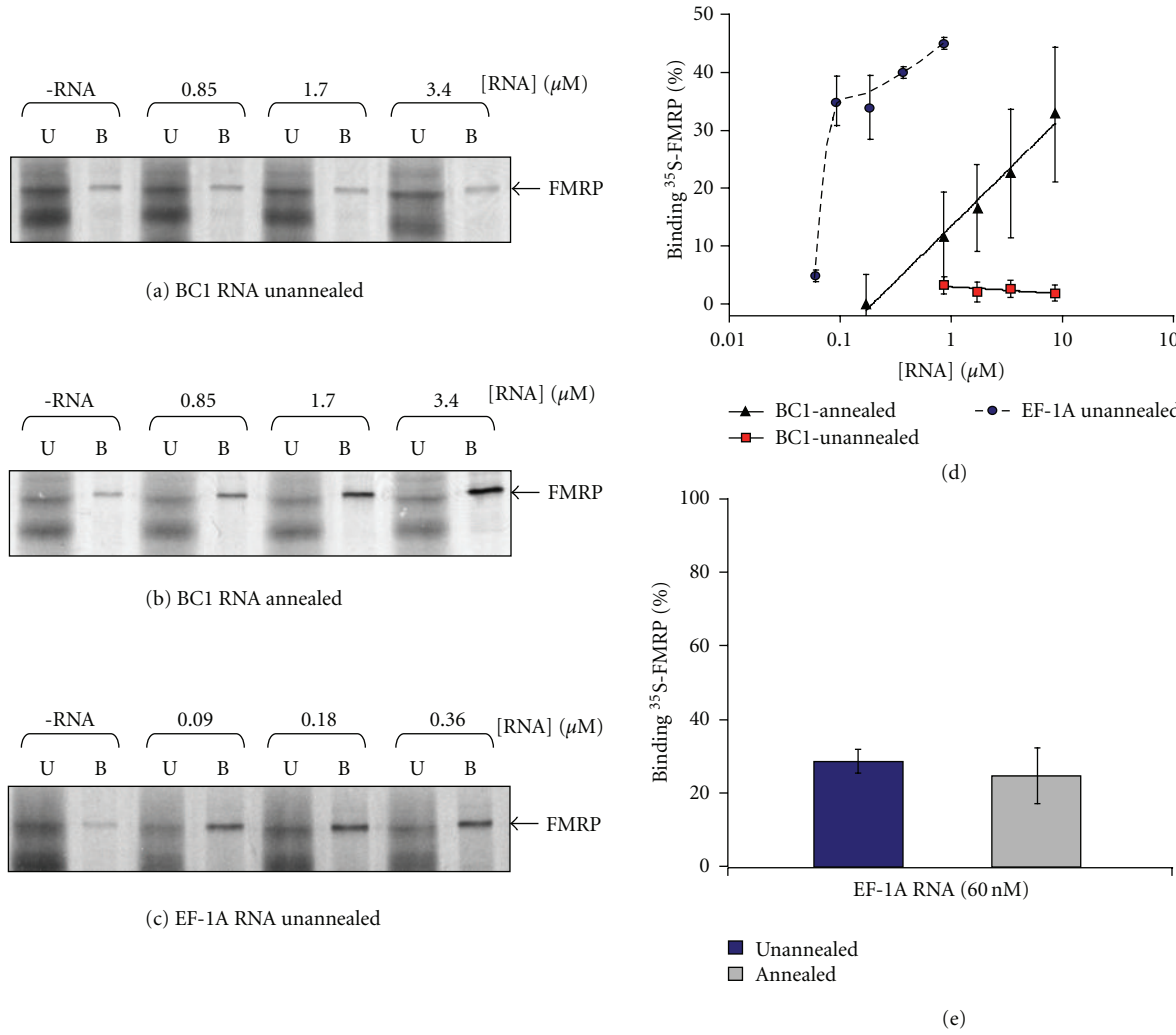


Figure 2: FMRP interacts weakly with annealed BC1 RNA. Equal amounts of ^{35}S-FMRP was titrated with various amounts of (a) unannealed biotinylated-BC1 RNA, (b) annealed biotinylated-BC1 RNA, and (c) unannealed biotinylated-eEF-1A RNA, as indicated. Bound (B) and unbound (U) ^{35}S-FMRP was recovered, resolved by SDS-PAGE and subject to autoradiography. Nonspecific binding to the SoftLink avidin resin is shown in the -RNA lanes. (d) Graphical analysis of the above data for 4 reactions per concentration. (e) Binding of annealed or unannealed eEF1A RNA (60 nM) to ^{35}S-FMRP. The percent binding corrected for background of 4 reactions per RNA type is plotted ($P = .45$ by ANOVA).

RNA conformer(s) or affect their resistance to RNase V1 (not shown).

These data unequivocally demonstrate that the conformer populations of annealed and unannealed BC1 RNA differ, and this difference is due to an increase in the amount of stable RNA duplexes in the annealed RNA. Moreover, the data support the hypothesis that the binding between FMRP and annealed BC1 RNA results from a unique conformer that is absent from the unannealed BC1 RNA conformer population.

3.4. Functional Dissection of the FMRP BC1 RNA Interaction. BC1 RNA contains several distinct higher order structural elements. Its 5′ end contains two cis-acting spatial targeting elements, DTE1 and DTE2 [45]. The former is required for somatic export of BC1 RNA into dendrites, while the latter specifies long-range distal delivery and is mediated by a prominent GA-type kink turn (KT) motif in the apical region of the 5′ BC1 domain [45]. The 3′ 60 bases of BC1 RNA contains an abbreviated A-rich region and a 3′-terminal stem-loop structure [65] that has been shown to bind synergistically to the eucaryotic initiation factor 4A (eIF4A) and the polyadenylation binding protein (PABP) [41]. To further probe this interaction, BC1 RNA was functionally dissected, and these two elements were examined individually. As shown in Figure 4(a), RNase VI treatment of the 5′ 75 bases of BC1 RNA recapitulated the results of full-length BC1 RNA indicating that it forms a stable secondary structure in the absence of the 3′ end. This agrees with modeling studies using the Zuker M-fold

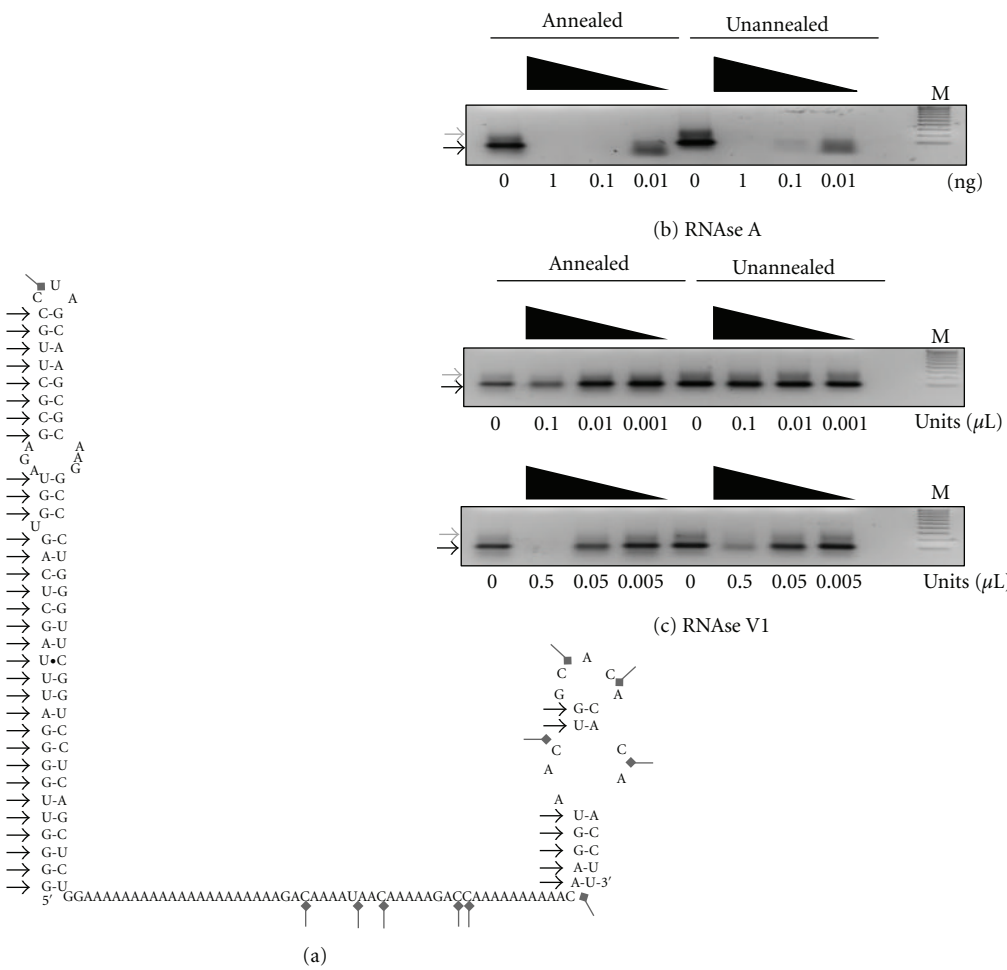


Figure 3: Annealed BC1 RNA and unannealed BC1 RNA differ structurally. (a) Secondary structure model of BC1 RNA from M-fold. Black arrows indicate potential RNase VI cleavage sites (both sides of the stem); gray arrows mark potential RNase A cleavage sites. (b) Serial treatment of annealed and unannealed full-length BC1 RNA with RNase A starting at 1 ng as indicated. (c) Ten-fold serial treatment of annealed and unannealed full-length BC1 RNA with RNase VI starting with 0.1 units/μL (upper panel), or 0.5 units/μL (lower panel) as indicated. RNA was visualized by ethidium bromide staining. Arrows mark the major conformers that can be resolved in this system. Lane M shows 100 bp molecular weight markers.

algorithm, which show that the first 75 bases of BC1 RNA fold identically to that portion of the full-length molecule, Figure 4(b). On the other hand, annealing had no effect on the sensitivity of the 3′ 60 bases of BC1 RNA towards RNase VI, Figure 5. Thus, these data demonstrate that the difference in the conformer populations of annealed and unannealed BC1 RNA most likely arise from alterations in the 5′ end of the molecule.

We next examined whether FMRP preferentially bound either of the dissected RNAs. Therefore, ^{35}S-FMRP was incubated with 2 μM annealed versions of each RNA; binding was subsequently assessed by affinity capture and compared to that of annealed full-length BC1 RNA. We found that the 5′ 75 bases of BC1 RNA bound slightly less, than full-length BC1 RNA, Figure 6(a); however, the difference did not rise to the level of statistical significance ($P = .15$, ANOVA). Truncation of this RNA by a 15 base deletion at its 3′ end further decreased binding ($P = .08$, ANOVA). Significantly, this 60 base RNA is not expected to recapitulate the folding of the first 60 bases of full-length BC1 RNA, Figure 6(b). On the other hand, the 3′ 60 bases of BC1 RNA evinced no evidence of binding, Figure 6(a). These data suggest that the major determinant that FMRP recognizes in annealed BC1 RNA is the hairpin structure of the 5′ end. To test this further, we also conducted binding studies using an 85 base RNA harboring the HIV1 TAR hairpin, Figure 6(c). The results show that this RNA binds with the same affinity as annealed full-length BC1 RNA.

3.5. Specificity and Requirements of the FMRP BC1 RNA Interaction. Different preparations or sources of FMRP has been posited as a potential explanation for the divergent BC1 RNA-binding data obtained by different laboratories [59, 60] and a likely cause of such differences is in the posttranslation modifications that occur in each. To begin to address this question, we examined whether differences in posttranslational arginine methylation could alter the binding affinity of annealed BC1 RNA and FMRP. Specifically, we compared the

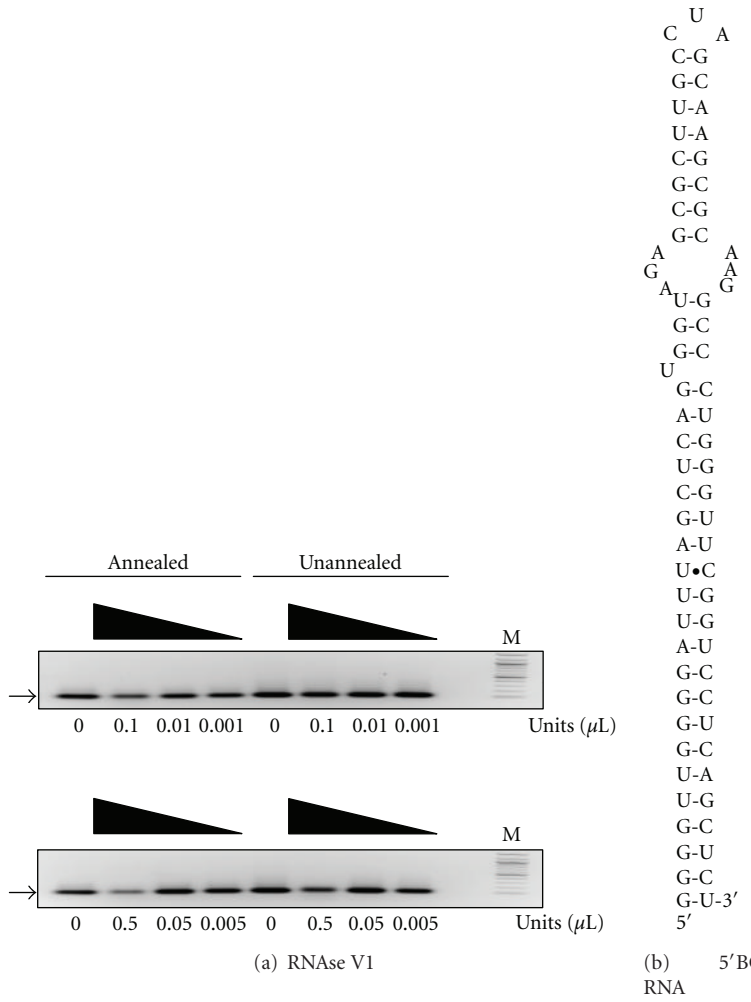


Figure 4: Functional dissection of structural elements in the 5′ end of BC1 RNA. (a) Ten-fold serial treatment of annealed and unannealed 5′BC1-75 RNA with RNase VI starting at 0.1 units/μL (upper panel) or 0.5 units/μL (lower panel). Arrow marks the major conformer. Lane M shows 100 bp molecular weight markers. (b) Secondary structure model of 5′BC1-75 RNA from M-fold.

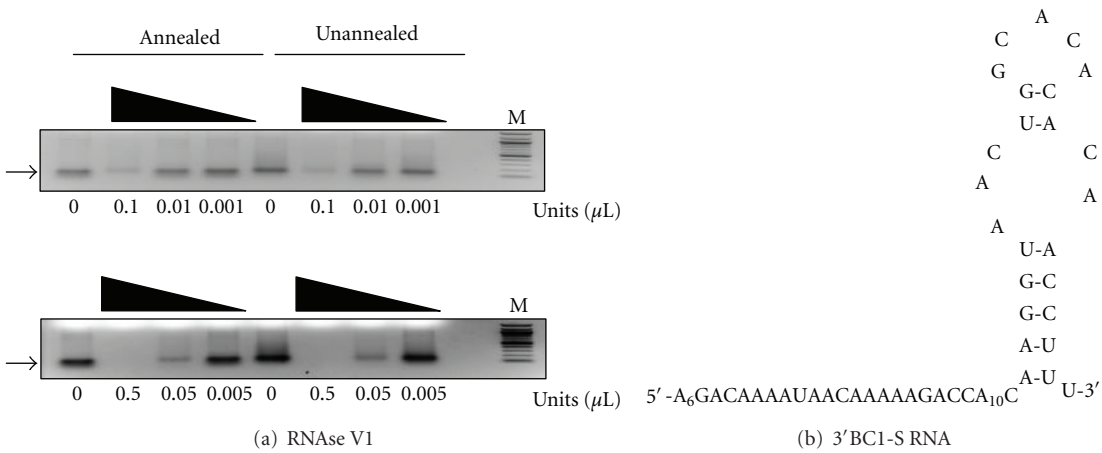


Figure 5: Functional dissection of structural elements in the 3′ end of BC1 RNA. (a) Ten-fold serial treatment of annealed and unannealed 3′BC1-60 RNA with RNase VI starting at 0.1 units/μL (upper panel) or 0.5 units/μL (lower panel). Arrow marks the major conformer. Lane M shows 100 bp molecular weight markers. (b) Secondary structure model of 3′BC1-60 RNA from M-fold.

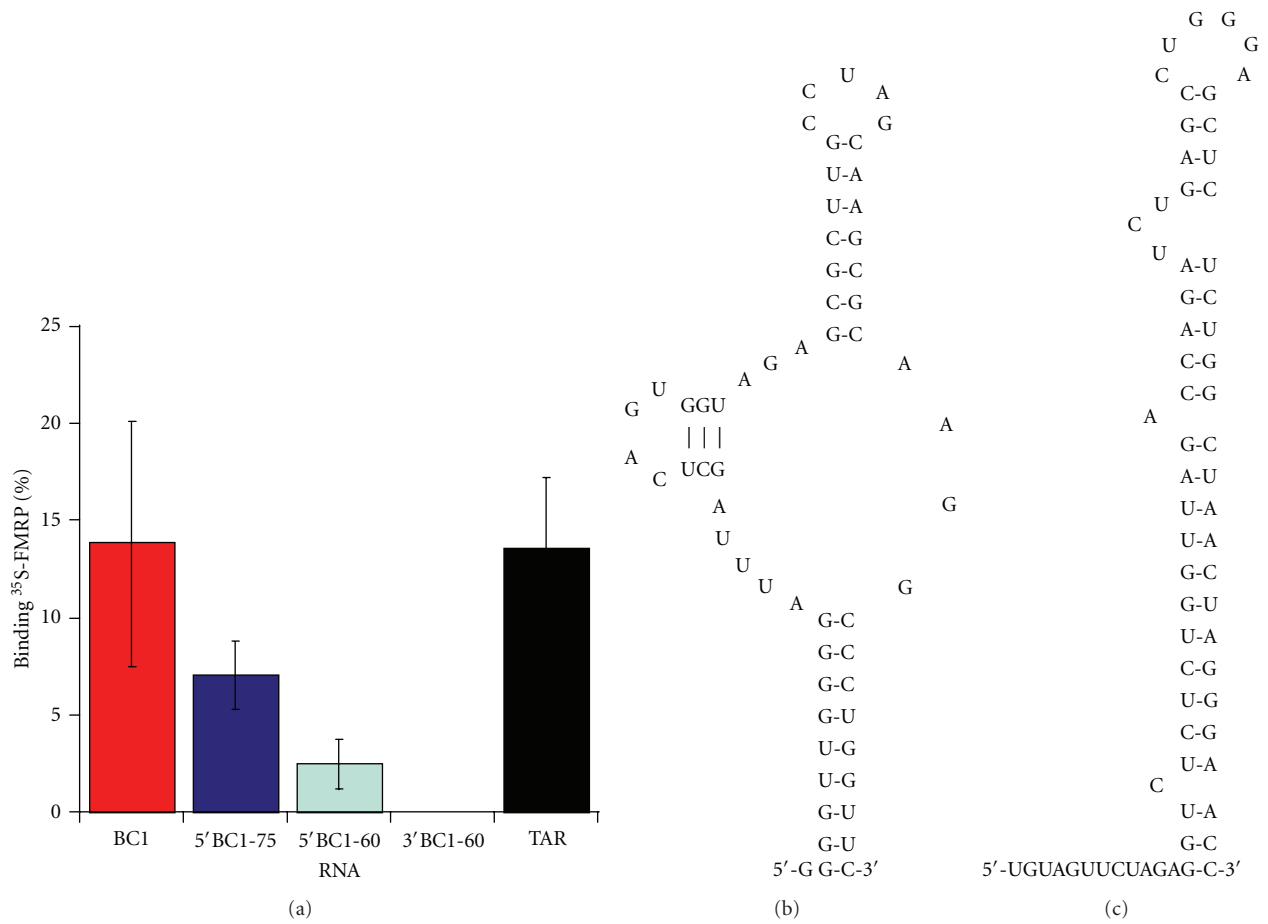


Figure 6: Interaction of FMRP with the 5′ and 3′ ends of BC1 RNA. (a) ^{35}S-FMRP was incubated with 2 μM annealed full-length BC1 RNA, or annealed versions of the first 75 b of BC1 RNA (5′BC1-75), the first 60 b of BC1 RNA (5'BC1-60) the last 60 b of BC1 RNA (3′BC1-60) and an 85 b transcript containing the HIV1 TAR hairpin. The percent binding corrected for background of 6 reactions per protein is plotted. (b) Secondary structure model of 5′BC1-60 RNA from M-fold. (c) Secondary structure model of the HIV1 TAR RNA from M-fold; the 27 b leader sequence has no effect upon the folding.

ability of full-length FMRP and the ability of an alternatively spliced variant $FMRP_{Ex15c}$ to bind 2 μM annealed BC1 RNA. Previous studies showed that full-length FMRP is readily methylated in its RG-rich region, while $FMRP_{15c}$, which lacks 25 amino acids upstream of the RG-rich region, is refractory to arginine methylation [66]. Figure 7(a) shows that binding of annealed BC1 RNA to $FMRP_{15c}$ was less than it was to full-length FMRP, but the difference did not rise to statistical significance.

We next examined whether the binding between annealed BC1 RNA and FMRP was unique. As an initial indicator, 2 μM annealed BC1 RNA was incubated with equimolar amounts of fragile X family members, ^{35}S-FMRP, ^{35}S-FXR1P and ^{35}S-FXR2P (full-length forms); subsequently, binding was assessed as previously described. As shown in Figure 7(a), both FXR1P and FXR2P bound as well as FMRP. To confirm and extend these results binding studies between annealed BC1 RNA and luciferase or eIF4A were performed. Luciferase does not contain any RNA-binding motifs and does not bind to poly (rA), poly (rG), or eEF-1A mRNA, Figure 7(b). Correspondingly, we found that luciferase did not interact with annealed BC1 RNA. On the other hand, the RNA helicase, eIF4A, bound weakly annealed BC1 RNA under these conditions, Figure 7(a). These data indicate that in addition to FMRP annealed BC1 RNA can interact interchangeably with at least three other RNA-binding proteins.

It has been previously proposed that BC1 RNA interacts with the N-terminal domain (NTD) of FMRP [40]. Nevertheless, Wang et al. did not observe an interaction between unannealed BC1 RNA and a 280 amino acid NTD construct, $FMRP_{1-280}$ [41]. To complete this analysis, we also assessed the interaction of annealed BC1 RNA with $FMRP_{1-280}$. As shown in Figure 7(a), annealed BC1 RNA did not bind to ^{35}S-$FMRP_{1-280}$. To determine whether ^{35}S-$FMRP_{1-280}$ lacked the ability to interact with RNA, we asked whether it could to bind to homoribopolymers. As shown in Figure 7(c), ^{35}S-$FMRP_{1-280}$ was unable to bind to poly(rA), poly(rG), or poly(rI : rC). These data suggest that ^{35}S-$FMRP_{1-280}$ does not significantly interact with ribonucleic acids under the conditions used. Interestingly, however, while a shorter NTD, that is, ^{35}S-$FMRP_{1-204}$, was also unable to bind to annealed

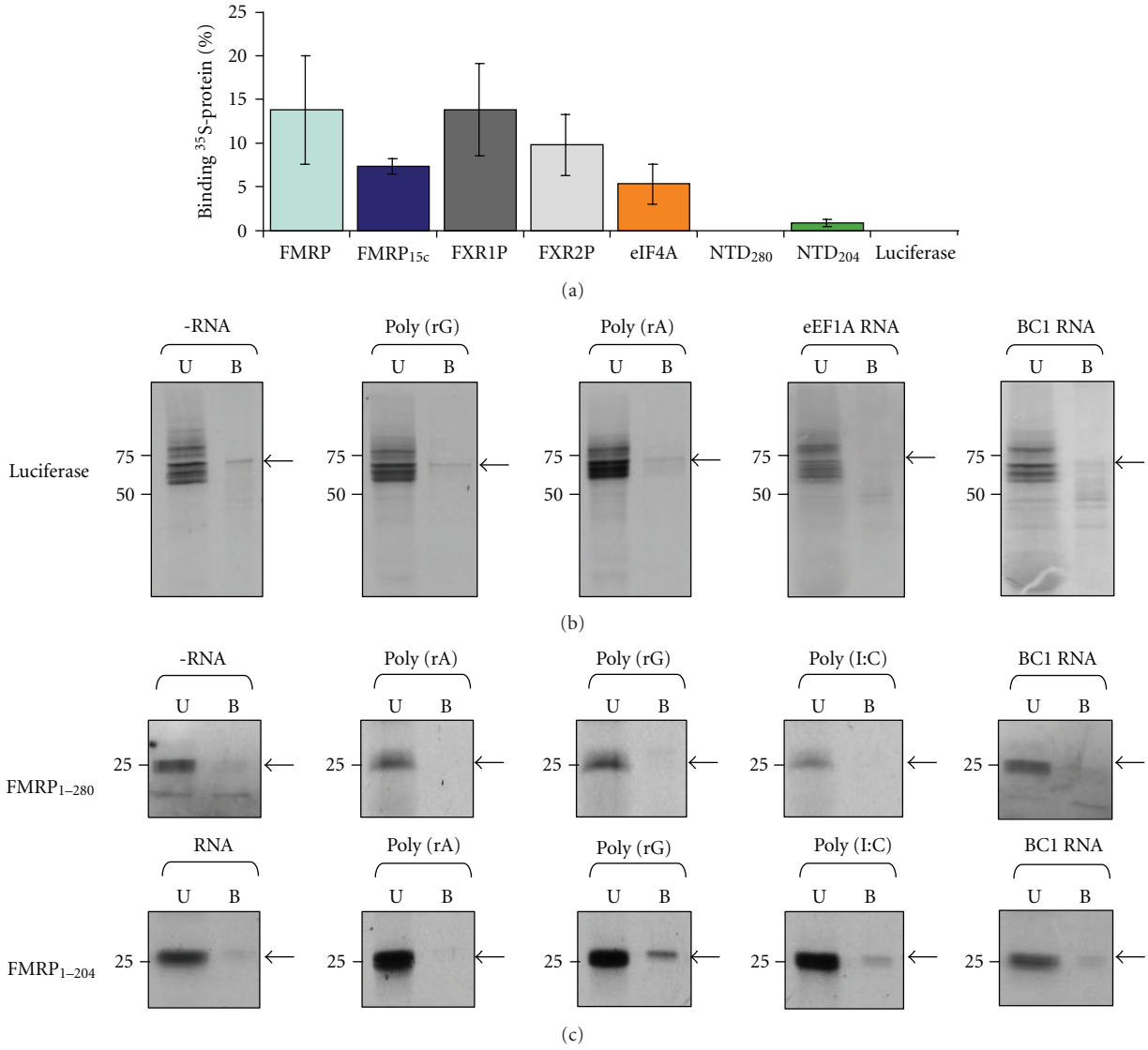


Figure 7: Annealed BC1 RNA binds pleiotropically to the FXRPs. (a) Equimolar amounts of ^{35}S-FMRP, ^{35}S-FMRP$_{15c}$,^{35}S-FXR1P, ^{35}S-FXR2P, ^{35}S-eIF4A, ^{35}S-FMRP$_{280}$, ^{35}S-FMRP$_{204}$, and ^{35}S-Luciferase were incubated with 2 μM annealed BC1 RNA. The percent binding corrected for background of 6 reactions per protein is plotted. The relative levels of FMRP and FMRP$_{15c}$, FMRP and FXR1P and FMRP and FXR2P were not significantly different ($P = .17$, $P = .67$ and $P = .27$, resp., ANOVA). However, the relative levels of FMRP and eIF4A and FMRP and NTD$_{204}$ were significantly different ($P = .038$ and $P = .007$, resp., ANOVA). Since no binding above background was observed for NTD$_{280}$ and Luciferase relative differences were not assessed. (b) Pull-down assays between ^{35}S-luciferase and poly (rG), poly (rA), eEF1A RNA (90 nM) and annealed BC1 RNA (5 μM) showing the unbound (U) and bound (B) fractions. Binding in the absence of RNA is shown for the negative control. (c) Pull-down assays between ^{35}S-NTD$_{280}$ or ^{35}S-NTD$_{204}$ and poly (rG), poly (rA), poly (rI : rC) and annealed BC1 RNA (5 μM) showing the unbound (U) and bound (B) fractions. Binding in the absence of RNA is shown for the negative control.

BC1 RNA, Figure 7(a), it was able to bind specifically to poly(rG) and to a lesser extent, poly(rI : rC), Figure 7(c). Thus, annealed BC1 RNA does not interact significantly with the 204 N-terminal residues of FMRP.

4. Conclusions

Disparate data obtained using electrophoretic mobility shift assays have been published concerning the binding of FMRP to BC1 RNA [36, 40, 47]. To investigate this discrepancy, we turned to an affinity capture assay to measure RNA binding to FMRP under the assumption that the results from different assays should converge to produce a common result. This particular assay was first described by Boehlens et al. and applied to the interactions of FMRP and nucleic acids by Ashley et al. [51, 67]. Here, we demonstrate that in 150 mM KCl and 125 mM NaCl, the assay has a dynamic range of RNA binding of at least two logs. Thus, it is able to distinguish high affinity binding from low affinity binding.

We specifically developed these assay conditions, which do not utilize tRNA as a general nonspecific binding competitor, because it had previously been demonstrated that both tRNA and BC1 RNA bind to FMRP in a low salt buffer with nearly identical dissociation constants (K_ds) [42], and because tRNA was shown to directly displace BC1 RNA from FMRP [47]. Thus, tRNA, while admittedly effective, may not be the best blocking reagent for examining FMRP's interaction with BC1 RNA. On the other hand, FMRP's association with polyribosomes [37, 38], translationally repressed ribosomes within neuronal granules [68], and its function in translational regulation [17, 19, 27] indicates that it operates in an environment with a relatively high local concentration of tRNA, which would ensure a direct competition with BC1 RNA. Indeed, previously published immunoprecipitation experiments, which are carried out in the presence of a large excess of tRNA (both added and from the endogenous tRNA present during cell lysis) and which fail to find BC1 RNA among the precipitated transcripts are entirely consistent with a direct competition between BC1 RNA and tRNA [47]. Here, we demonstrate that BC1 RNA produced simply by in vitro transcription does not interact with FMRP even in the absence of tRNA, confirming our previous results using different assays and different binding conditions [41]. Nevertheless, by artificially annealing the BC1 RNA transcript, we found that binding occurred although it was at least a hundred-fold weaker than that between FMRP and eEF-1A mRNA. Previous protocols examining the interaction between BC1 RNA and FMRP have not indicated that the RNA was pretreated or annealed before binding to FMRP was initiated [36, 40, 47]; therefore, it must be assumed that it was not, and hence, this treatment is unique to this work.

To try to understand the result, we first determined whether enhanced binding due to annealing was a general feature of FMRP's interaction with RNA. A search of the current literature produced mixed results. Darnell et al., working with kissing complex RNA a small double hairpin with additional loop-loop tertiary interactions that binds to FMRP's KH_2 domain, preannealed the RNA before binding was initiated [30]. Likewise, Ramos et al. and Zanotti et al. preannealed the synthetic G-rich RNA, sc1, to form G-quartets before binding to RGG box peptides [69, 70]. Finally, Bechara et al. renatured SoSLIP RNA, a small multiple hairpin-containing RNA, before they used it in binding reactions with FMRP [19]. In contrast, longer RNAs containing U-rich motifs such as MBP mRNA [71], the FMR1 3′UTR [21], eEF-1A mRNA [17], and BMP receptor mRNA [41] do not require preannealing to bind tightly to FMRP. However, none of these studies directly compare binding between annealed and unannealed RNAs. Therefore, we examined the effect annealing eEF-1A mRNA had on its binding to FMRP. Under conditions where FMRP's interaction with unannealed eEF-1A was linearly related to the input RNA we found that preannealing the RNA had no measurable effect upon the binding. Thus, while not a general requirement for FMRP binding preannealing may be necessary for smaller RNAs whose structure is more susceptible to changes in temperature, buffer pH, concentration and type of ions and RNA concentration [70].

We next determined whether annealed BC1 RNA was structurally altered from unannealed BC1 RNA. Indeed, we showed that the annealed form exhibited an increased susceptibility to cleavage by RNase VI. These data support the hypothesis that a particular BC1 RNA conformer; that is, one containing a unique doubled-stranded RNA structure was the true FMRP interactor. Functional dissection of BC1 RNA into either a 5′ ID element or a 3′ element demonstrated that the major determinant that FMRP binds to is located in the 5′ 75 bases of the molecule. As FMRP also binds to tRNA and TAR RNA, these data imply that FMRP can weakly recognize hairpin-containing RNAs.

Another difference between the published FMRP BC1 RNA interaction studies was that one group used full-length recombinant FMRP produced from baculovirus [36, 40], while the other group used full-length recombinant FMRP produced in *E. coli* [47]. A potential concern here is that each of these preparations may be differentially posttranslationally modified. Indeed, it is well known that FMRP is subject to posttranslational methylation of arginine residues in its RG-rich region [18, 31]. However, while protein arginine methyltransferases are present in *Spodoptera frugiperda* [72], the host cells used in baculovirus FMRP production [40], they are absent from *E. coli* [73, 74]. Thus, it is possible that the presence or absence of methyl-arginine residues might directly or indirectly affect an interaction between FMRP and BC1 RNA. In the present work, we opted to produce full-length FMRP in rabbit reticulocyte lysates (RRL), which generates partially methylated FMRP [66]. Again, we assumed that posttranslational modifications arising from source differences could be a modulating but not *the* determining factor in a putative FMRP BC1 RNA interaction. Precedence for this assumption may be seen in the work of Stetler et al. who showed that methylating FMRP decreased, but did not abrogate, the direct binding of two G-quartet RNAs to FMRP's RG-rich region [56]. Consistent with the work using *E. coli*-purified FMRP [47] and with the work of Wang et al. who used unmethylated FMRP produced in cell-free wheat germ lysates to examine BC1 RNA binding [41], we observed minimal interactions between RRL-produced FMRP and unannealed BC1 RNA. Thus, methylating FMRP's arginine residues does not enhance its affinity for unannealed BC1 RNA. Furthermore, although binding between annealed BC1 RNA and the methylation refractory form of FMRP ($FMRP_{15c}$) was reduced compared to methylated FMRP, it was not a statistically significant reduction. Thus, methylarginine residues in FMRP's RG-rich region *are not absolutely required* for this interaction either.

We also probed the exclusivity of FMRP's interaction with annealed BC1 RNA. Here, we found that interactions of similar magnitude occur with fragile X paralogs FXR1P and FXR2P and to a lesser extent, eIF4A. These results demonstrate that BC1 RNA binds indiscriminately to RNA-binding proteins. Interestingly, contrasting results were recently obtained for kissing complex (kc) RNA [34], where it was shown that this RNA exclusively interacts with the

fragile X paralogs but not with other related KH domain-containing RNA-binding proteins.

As all three paralogs have highly homologous N-terminal sequences but are more divergent in their C-terminal ends [75], we inquired whether annealed BC1 RNA interacted with specific N-terminal fragments of FMRP (NTDs). We did not observe an interaction between annealed BC1 RNA and $FMRP_{1-280}$, a construct comprising both N-terminal residues and the first KH domain (KH_1). The results were consistent with previous work showing that this construct did not interact with unannealed BC1 RNA [41]. However, as $FMRP_{1-280}$ did not interact with other nucleic acids under these conditions, we could not rule out the possibility that the protein was mis-folded although previous physical evidence suggests this is not the case [50]. Therefore, we also examined the interaction between annealed BC1 RNA and $FMRP_{1-204}$, a construct that does interact with nucleic acids and is expected "to be at least partially folded and be monodisperse" under the conditions it was used [50]. Nevertheless, annealed BC1 RNA did not interact with this FMRP fragment either. Hence, the first 204 residues of FMRP, which includes two Tudor domains and an alpha helix that is essential for domain stability [76], are not required for the interaction of annealed BC1 RNA with FMRP. Using different conditions, Zalfa et al. have suggested that BC1 RNA binds to an NTD comprising amino acids 1–217 [40]. The additional 13 residues of this construct comprise part of another alpha helix, which stabilizes the folding of FMRP's KH_1 domain [50]. While it is possible that amino acids 205–217 may confer binding specificity to annealed BC1 RNA, Zalfa et al. showed that a construct comprising this alpha helix along with the KH_1 domain (residues 205–280) did not interact with BC1 RNA [40]. However, additional studies using other constructs will be needed to fully address the question of which FMRP residues are necessary and sufficient for its interaction with annealed BC1 RNA.

Although we demonstrate an in vitro interaction between FMRP and BC1 RNA, its nature and its physiological significance remain elusive. For example, the structure of the unique BC1 RNA conformer that FMRP binds to has not been defined other than the fact that it contains double-stranded RNA and is primarily found in the 5′ half of the molecule. In fact, it is possible based on the rather high RNA concentrations used in the annealing reaction that FMRP may be interacting with a dimer of BC1 RNA. This would be analogous to the formation of intermolecular G-quartet RNAs such as MAP1B RNA [69]. Clearly, a more comprehensive biophysical analysis of the parameters of this interaction is needed to fully address this issue. Regardless, the ultimate significance of this observation is questionable given that the formation of this conformer does not occur at temperatures that mammalian cells can survive. Although one might postulate that a chaperone protein may be able to mitigate the temperature requirement for annealing [77, 78], it is clear that FMRP cannot be this chaperone, because it is unable convert unannealed BC1 RNA into a molecule that it binds under physiological conditions. Given this, our data generated using different constructs, different preparations and different methods, converge with the published work of Iacoangeli et al. [47] to support a model in which FMRP and BC1 RNA operate independently of each other to control protein synthesis in neuronal processes.

Abbreviations

AGESA: Agarose electrophoretic mobility shift assays
BC1 RNA: Brain cytoplasmic RNA
eEF-1A: Eucaryotic elongation factor 1A
eIF4A: Eucaryotic initiation factor 4A
EMSA: Electrophoretic mobility shift assays
FMRP: Fragile X mental retardation protein
NTD: N-terminal domain
PABP: Polyadenylation binding protein
RBP: RNA binding protein
RRL: Rabbit reticulocyte lysate.

Acknowledgments

The authors thank Drs. Henri Tiedge, Edward Khandjian, Stefan Kindler and Jurgen Brosius for critically reading this paper. This work was supported by the New York State Research Foundation for Mental Hygiene.

References

[1] C. Bagni and W. T. Greenough, "From mRNP trafficking to spine dysmorphogenesis: the roots of fragile X syndrome," *Nature Reviews Neuroscience*, vol. 6, no. 5, pp. 376–387, 2005.

[2] B. Bardoni, L. Davidovic, M. Bensaid, and E. W. Khandjian, "The fragile X syndrome: exploring its molecular basis and seeking a treatment," *Expert Reviews in Molecular Medicine*, vol. 8, no. 8, pp. 1–16, 2006.

[3] P. Jin and S. T. Warren, "New insights into fragile X syndrome: from molecules to neurobehaviors," *Trends in Biochemical Sciences*, vol. 28, no. 3, pp. 152–158, 2003.

[4] W. T. O'Donnell and S. T. Warren, "A decade of molecular studies of fragile X syndrome," *Annual Review of Neuroscience*, vol. 25, pp. 315–338, 2002.

[5] V. Mercaldo, G. Descalzi, and M. Zhuo, "Fragile X mental retardation protein in learning-related synaptic plasticity," *Molecules and Cells*, vol. 28, no. 6, pp. 501–507, 2009.

[6] S. M. Till, "The developmental roles of FMRP," *Biochemical Society Transactions*, vol. 38, no. 2, pp. 507–510, 2010.

[7] A. Lo-Castro, E. D'Agati, and P. Curatolo, "ADHD andgenetic syndromes," *Brain and Development*. In press.

[8] P. Hagerman and C. E. Stafstrom, "Origins of epilepsy in fragile X syndrome," *Epilepsy Currents*, vol. 9, no. 4, pp. 108–112, 2009.

[9] B. J. Flynn, S. M. Myers, P. J. Cera, and J. J. Mowad, "Testicular torsion in an adolescent with fragile X syndrome," *Pediatrics*, vol. 109, no. 1, p. E16, 2002.

[10] L. Chakrabarti and K. E. Davies, "Fragile X syndrome," *Current Opinion in Neurology*, vol. 10, no. 2, pp. 142–147, 1997.

[11] B. A. Oostra and R. Willemsen, "A fragile balance: FMR1 expression levels," *Human Molecular Genetics*, vol. 12, no. 2, pp. R249–R257, 2003.

[12] P. J. Hagerman and R. J. Hagerman, "Fragile X-associated, tremor/ataxia syndrome (FXTAS)," *Mental Retardation and Developmental Disabilities Research Reviews*, vol. 10, no. 1, pp. 25–30, 2004.

[13] H. Tan, H. Li, and P. Jin, "RNA-mediated pathogenesis in fragile X-associated disorders," *Neuroscience Letters*, vol. 466, no. 2, pp. 103–108, 2009.

[14] B. A. Oostra and R. Willemsen, "FMR1: a gene with three faces," *Biochimica et Biophysica Acta*, vol. 1790, no. 6, pp. 467–477, 2009.

[15] H. Siomi, M. C. Siomi, R. L. Nussbaum, and G. Dreyfuss, "The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein," *Cell*, vol. 74, no. 2, pp. 291–298, 1993.

[16] B. Laggerbauer, D. Ostareck, E.-M. Keidel, A. Ostareck-Lederer, and U. Fischer, "Evidence that fragile X mental retardation protein is a negative regulator of translation," *Human Molecular Genetics*, vol. 10, no. 4, pp. 329–338, 2001.

[17] Y. J. Sung, N. Dolzhanskaya, S. L. Nolin, T. Brown, J. R. Currie, and R. B. Denman, "The fragile X mental retardation protein FMRP binds elongation factor 1A mRNA and negatively regulates its translation in vivo," *Journal of Biological Chemistry*, vol. 278, no. 18, pp. 15669–15678, 2003.

[18] R. B. Denman, N. Dolzhanskaya, and Y. J. Sung, "Regulating a translational regulator: mechanisms cells use to control the activity of the fragile X mental retardation protein," *Cellular and Molecular Life Sciences*, vol. 61, no. 14, pp. 1714–1728, 2004.

[19] E. G. Bechara, M. C. Didiot, M. Melko et al., "A novel function for fragile X mental retardation protein in translational activation," *PLoS Biology*, vol. 7, no. 1, article e16, 2009.

[20] M. Fahling, R. Mrowka, A. Steege et al., "Translational regulation of the human achaete-scute homologue-1 (hASH1) by fragile x-mental retardation protein (FMRP)," *Journal of Biological Chemistry*, vol. 284, no. 7, pp. 4255–4266, 2009.

[21] Y. J. Sung, J. Conti, J. R. Currie, W. T. Brown, and R. B. Denman, "RNAs that interact with the fragile X syndrome RNA binding protein FMRP," *Biochemical and Biophysical Research Communications*, vol. 275, no. 3, pp. 973–980, 2000.

[22] J. C. Darnell, K. B. Jensen, P. Jin, V. Brown, S. T. Warren, and R. B. Darnell, "Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function," *Cell*, vol. 107, no. 4, pp. 489–499, 2001.

[23] L. Chen, S.-W. Yun, J. Seto, W. Liu, and M. Toth, "The fragile X mental retardation protein binds and regulates a novel class of mRNAs containing u rich target sequences," *Neuroscience*, vol. 120, no. 4, pp. 1005–1017, 2003.

[24] V. Brown, P. Jin, S. Ceman et al., "Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome," *Cell*, vol. 107, no. 4, pp. 477–487, 2001.

[25] K. Y. Miyashiro, A. Beckel-Mitchener, T. P. Purk et al., "RNA cargoes associating with FMRP reveal deficits in cellular functioning in Fmr1 null mice," *Neuron*, vol. 37, no. 3, pp. 417–431, 2003.

[26] C. Schaeffer, B. Bardoni, J.-L. Mandel, B. Ehresmann, C. Ehresmann, and H. Moine, "The fragile X mental retardation protein binds specifically to its mRNA *via* a purine quartet motif," *EMBO Journal*, vol. 20, no. 17, pp. 4803–4813, 2001.

[27] S. Ceman, W. T. O'Donnell, M. Reed, S. Patton, J. Pohl, and S. T. Warren, "Phosphorylation influences the translation state of FMRP-associated polyribosomes," *Human Molecular Genetics*, vol. 12, no. 24, pp. 3295–3305, 2003.

[28] P. Jin, D. C. Zarnescu, S. Ceman et al., "Biochemical and genetic interaction between the fragile X mental retardation protein 3nd the microRNA pathway," *Nature Neuroscience*, vol. 7, no. 2, pp. 113–117, 2004.

[29] L. N. Antar, R. Afroz, J. B. Dictenberg, R. C. Carroll, and G. J. Bassell, "Metabotropic glutamate receptor activation regulates fragile X mental retardation protein and Fmr1 mRNA localization differentially in dendrites and at synapses," *Journal of Neuroscience*, vol. 24, no. 11, pp. 2648–2655, 2004.

[30] J. C. Darnell, C. E. Fraser, O. Mostovetsky et al., "Kissing complex RNAs mediate interaction between the fragile-X mental retardation protein KH2 domain and brain polyribosomes," *Genes and Development*, vol. 19, no. 8, pp. 903–918, 2005.

[31] N. Dolzhanskaya, G. Merz, J. M. Aletta, and R. B. Denman, "Methylation regulates the intracellular protein-protein and protein-RNA interactions of FMRP," *Journal of Cell Science*, vol. 119, no. 9, pp. 1933–1946, 2006.

[32] A. Aschrafi, B. A. Cunningham, G. M. Edelman, and P. W. Vanderklish, "The fragile X mental retardation protein and group I metabotropic glutamate receptors regulate levels of mRNA granules in brain," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 6, pp. 2180–2185, 2005.

[33] L. Hou, M. D. Antion, D. Hu, C. M. Spencer, R. Paylor, and E. Klann, "Dynamic translational and proteasomal regulation of fragile X mental retardation protein controls mGluR-dependent long-term depression," *Neuron*, vol. 51, no. 4, pp. 441–454, 2006.

[34] J. C. Darnell, C. E. Fraser, O. Mostovetsky, and R. B. Darnell, "Discrimination of common and unique RNA-binding activities among fragile X mental retardation protein paralogs," *Human Molecular Genetics*, vol. 18, no. 17, pp. 3164–3177, 2009.

[35] M. Melko and B. Bardoni, "The role of G-quadruplex in RNA metabolism: involvement of FMRP and FMR2P," *Biochimie*, vol. 92, no. 8, pp. 919–926, 2010.

[36] F. Zalfa, M. Giorgi, B. Primerano et al., "The fragile X syndrome protein FMRP associates with BC1 RNA and regulates the translation of specific mRNAs at synapses," *Cell*, vol. 112, no. 3, pp. 317–327, 2003.

[37] E. W. Khandjian, M.-E. Huot, S. Tremblay, L. Davidovic, R. Mazroui, and B. Bardoni, "Biochemical evidence for the association of fragile X mental retardation protein with brain polyribosomal ribonucleoparticles," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 36, pp. 13357–13362, 2004.

[38] G. Stefani, C. E. Fraser, J. C. Darnell, and R. B. Darnell, "Fragile X mental retardation protein is associated with translating polyribosomes in neuronal cells," *Journal of Neuroscience*, vol. 24, no. 33, pp. 7272–7276, 2004.

[39] L. Davidovic, M.-E. Huot, and E. W. Khandjian, "Lost once, the fragile X mental retardation protein is now back onto brain polyribosomes," *RNA Biology*, vol. 2, no. 1, pp. e1–e3, 2005.

[40] F. Zalfa, S. Adinolfi, I. Napoli et al., "Fragile X mental retardation protein (FMRP) binds specifically to the brain cytoplasmic RNAs BC1/BC200 via a novel RNA-binding motif," *Journal of Biological Chemistry*, vol. 280, no. 39, pp. 33403–33410, 2005.

[41] H. Wang, A. Iacoangeli, D. Lin et al., "Dendritic BC1 RNA in translational control mechanisms," *Journal of Cell Biology*, vol. 171, no. 5, pp. 811–821, 2005.

[42] C. Gabus, R. Mazroui, S. Tremblay, E. W. Khandjian, and J.-L. Darlix, "The fragile X mental retardation protein has nucleic

acid chaperone properties," *Nucleic Acids Research*, vol. 32, no. 7, pp. 2129–2137, 2004.

[43] H. Wang, A. Iacoangeli, S. Popp et al., "Dendritic BC1 RNA: functional role in regulation of translation initiation," *Journal of Neuroscience*, vol. 22, no. 23, pp. 10232–10241, 2002.

[44] H. Wang and H. Tiedge, "Translational control at the synapse," *Neuroscientist*, vol. 10, no. 5, pp. 456–466, 2004.

[45] I. A. Muslimov, A. Iacoangeli, J. Brosius, and H. Tiedge, "Spatial codes in dendritic BC1 RNA," *Journal of Cell Biology*, vol. 175, no. 3, pp. 427–439, 2006.

[46] F. Zalfa and C. Bagni, "Another view of the role of FMRP in translational regulation," *Cellular and Molecular Life Sciences*, vol. 62, no. 2, pp. 251–252, 2005.

[47] A. Iacoangeli, T. S. Rozhdestvensky, N. Dolzhanskaya et al., "On BC1 RNA and the fragile X mental retardation protein," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 2, pp. 734–739, 2008.

[48] J. G. Cheng, H. Tiedge, and J. Brosius, "Identification and characterization of BC1 RNP particles," *DNA and Cell Biology*, vol. 15, no. 7, pp. 549–559, 1996.

[49] V. Brown, K. Small, L. Lakkis et al., "Purified recombinant Fmrp exhibits selective RNA binding as an intrinsic property of the fragile X mental retardation protein," *Journal of Biological Chemistry*, vol. 273, no. 25, pp. 15521–15527, 1998.

[50] S. Adinolfi, C. Bagni, G. Musco, T. Gibson, L. Mazzarella, and A. Pastore, "Dissecting FMR1, the protein responsible for fragile X syndrome, in its structural and functional domains," *RNA*, vol. 5, no. 9, pp. 1248–1258, 1999.

[51] C. T. Ashley Jr., K. D. Wilkinson, D. Reines, and S. T. Warren, "FMR1 protein: conserved RNP family domains and selective RNA binding," *Science*, vol. 262, no. 5133, pp. 563–566, 1993.

[52] D. K. Price, F. Zhang, C. T. Ashley Jr., and S. T. Warren, "The chicken FMR1 gene is highly conserved with a CCT 5′-untranslated repeat and encodes an RNA-binding protein," *Genomics*, vol. 31, no. 1, pp. 3–12, 1996.

[53] Z. Li, Y. Zhang, L. Ku, K. D. Wilkinson, S. T. Warren, and Y. Feng, "The fragile X mental retardation protein inhibits translation via interacting with mRNA," *Nucleic Acids Research*, vol. 29, no. 11, pp. 2276–2283, 2001.

[54] M.-C. Didiot, Z. Tian, C. Schaeffer, M. Subramanian, J.-L. Mandel, and H. Moine, "The G-quartet containing FMRP binding site in FMR1 mRNA is a potent exonic splicing enhancer," *Nucleic Acids Research*, vol. 36, no. 15, pp. 4902–4912, 2008.

[55] R. B. Denman, "mRNPs take shape by CLIPPING and PAIRING," *BioEssays*, vol. 28, no. 11, pp. 1132–1143, 2006.

[56] A. Stetler, C. Winograd, J. Sayegh et al., "Identification and characterization of the methyl arginines in the fragile X mental retardation protein Fmrp," *Human Molecular Genetics*, vol. 15, no. 1, pp. 87–96, 2006.

[57] R. B. Denman and Y. J. Sung, "Species-specific and isoform-specific RNA binding of human and mouse fragile X mental retardation proteins," *Biochemical and Biophysical Research Communications*, vol. 292, no. 4, pp. 1063–1069, 2002.

[58] W. Xie, N. Dolzhanskaya, G. LaFauci, C. Dobkin, and R. B. Denman, "Tissue and developmental regulation of fragile X mental retardation 1 exon 12 and 15 isoforms," *Neurobiology of Disease*, vol. 35, no. 1, pp. 52–62, 2009.

[59] F. Zalfa, T. Achsel, and C. Bagni, "mRNPs, polysomes or granules: FMRP in neuronal protein synthesis," *Current Opinion in Neurobiology*, vol. 16, no. 3, pp. 265–269, 2006.

[60] C. Bagni, "On BC1 RNA and the fragile X mental retardation protein," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 17, p. E19, 2008.

[61] E. Bechara, L. Davidovic, M. Melko et al., "Fragile X related protein 1 isoforms differentially modulate the affinity of fragile X mental retardation protein for G-quartet RNA structure," *Nucleic Acids Research*, vol. 35, no. 1, pp. 299–306, 2007.

[62] L. Menon and M. Mihailescu, "Interactions of the G quartet forming semaphorin 3F RNA with the RGG box domain of the fragile X protein family," *Nucleic Acids Research*, vol. 35, no. 16, pp. 5379–5392, 2007.

[63] K. Zou, J. Liu, N. Zhu et al., "Identification of FMRP-associated mRNAs using yeast three-hybrid system," *American Journal of Medical Genetics*, vol. 147, no. 6, pp. 769–777, 2008.

[64] R. B. Denman, "The FMRP-BC1 RNA controversy: tempest in a teapot or symptomatic of a larger problem," in *The Molecular Basis of Fragile X Syndrome*, Y. J. Sung and R. B. Denman, Eds., pp. 65–79, Research Signpost, Trivandrum, India, 2005.

[65] T. S. Rozhdestvensky, A. M. Kopylov, J. Brosius, and A. Huttenhofer, "Neuronal BC1 RNA structure: evolutionary conversion of a tRNA(Ala) domain into an extended stem-loop structure," *RNA*, vol. 7, no. 5, pp. 722–730, 2001.

[66] N. Dolzhanskaya, G. Merz, and R. B. Denman, "Alternative splicing modulates protein arginine methyltransferase-dependent methylation of fragile X syndrome mental retardation protein," *Biochemistry*, vol. 45, no. 34, pp. 10385–10393, 2006.

[67] W. C. Boelens, E. J. R. Jansen, W. J. van Venrooij, R. Stripecke, I. W. Mattal, and S. I. Gunderson, "The human U1 snRNP-specific U1A protein inhibits polyadenylation of its own pre-mRNA," *Cell*, vol. 72, no. 6, pp. 881–892, 1993.

[68] M. A. Kiebler and G. J. Bassell, "Neuronal RNA granules: movers and makers," *Neuron*, vol. 51, no. 6, pp. 685–690, 2006.

[69] A. Ramos, D. Hollingworth, and A. Pastore, "G-quartet-dependent recognition between the FMRP RGG box and RNA," *RNA*, vol. 9, no. 10, pp. 1198–1207, 2003.

[70] K. J. Zanotti, P. E. Lackey, G. L. Evans, and M.-R. Mihailescu, "Thermodynamics of the fragile X mental retardation protein RGG box interactions with G quartet forming RNA," *Biochemistry*, vol. 45, no. 27, pp. 8319–8330, 2006.

[71] H. Wang, L. Ku, D. J. Osterhout et al., "Developmentally-programmed FMRP expression in oligodendrocytes: a potential role of FMRP in regulating translation in oligodendroglia progenitors," *Human Molecular Genetics*, vol. 13, no. 1, pp. 79–89, 2004.

[72] R. B. Denman, "Protein methyltransferase activities in commercial in vitro translation systems," *Journal of Biochemistry*, vol. 144, no. 2, pp. 223–233, 2008.

[73] J. D. Gary and S. Clarke, "RNA and protein interactions modulated by protein arginine methylation," *Progress in Nucleic Acid Research and Molecular Biology*, vol. 61, pp. 65–131, 1998.

[74] C. D. Krause, Z.-H. Yang, Y.-S. Kim, J.-H. Lee, J. R. Cook, and S. Pestka, "Protein arginine methyltransferases: evolution and assessment of their pharmacological and therapeutic potential," *Pharmacology and Therapeutics*, vol. 113, no. 1, pp. 50–87, 2007.

[75] E. W. Khandjian, "Biology of the fragile X mental retardation protein, an RNA-binding protein," *Biochemistry and Cell Biology*, vol. 77, no. 4, pp. 331–342, 1999.

[76] A. Ramos, D. Hollingworth, S. Adinolfi et al., "The structure of the N-terminal domain of the fragile X mental retardation protein: a platform for protein-protein interaction," *Structure*, vol. 14, no. 1, pp. 21–31, 2006.

[77] D. Herschlag, "RNA chaperones and the RNA folding problem," *Journal of Biological Chemistry*, vol. 270, no. 36, pp. 20871–20874, 1995.

[78] H. De Rocquigny, C. Gabus, A. Vincent, M.-C. Fournié-Zaluski, B. Roques, and J.-L. Darlix, "Viral RNA annealing activities of human immunodeficiency virus type 1 nucleocapsid protein require only peptide domains outside the zinc fingers," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 14, pp. 6472–6476, 1992.

SAGE-Hindawi Access to Research
Journal of Nucleic Acids
Volume 2011, Article ID 408053, 17 pages
doi:10.4061/2011/408053

Review Article

Use of Specific Chemical Reagents for Detection of Modified Nucleotides in RNA

Isabelle Behm-Ansmant,[1] Mark Helm,[2] and Yuri Motorin[1]

[1] *Laboratoire ARN-RNP Maturation-Structure-Fonction, Enzymologie Moléculaire et Structurale (AREMS), UMR 7214 CNRS-UHP, Nancy Université, boulevard des Aiguillettes, BP 70239, 54506 Vandoeuvre-les-Nancy, France*
[2] *Institute of Pharmacy and Biochemistry, Johannes Gutenberg-University of Mainz, Staudingerweg 5, 55099 Mainz, Germany*

Correspondence should be addressed to Yuri Motorin, iouri.motorine@maem.uhp-nancy.fr

Received 15 October 2010; Accepted 24 January 2011

Academic Editor: Dmitry A. Stetsenko

Naturally occurring cellular RNAs contain an impressive number of chemically distinct modified residues which appear posttranscriptionally, as a result of specific action of the corresponding RNA modification enzymes. Over 100 different chemical modifications have been identified and characterized up to now. Identification of the chemical nature and exact position of these modifications is typically based on 2D-TLC analysis of nucleotide digests, on HPLC coupled with mass spectrometry, or on the use of primer extension by reverse transcriptase. However, many modified nucleotides are silent in reverse transcription, since the presence of additional chemical groups frequently does not change base-pairing properties. In this paper, we give a summary of various chemical approaches exploiting the specific reactivity of modified nucleotides in RNA for their detection.

1. Introduction

Native cellular RNAs contain numerous modified residues resulting from specific action of various RNA modification enzymes. These RNA modifications are ubiquitous in nature, but the specific modification profile varies depending on the organism. Over 100 chemically distinct modified nucleotides have been identified so far mostly in tRNAs, rRNAs, snRNAs and some snoRNAs. From the chemical point of view, these modifications are highly diverse and almost any position of the nucleobases as well as the 2′-OH of the ribose has been found to be a target of modification enzymes (see Table 1 and below) [1–3].

Identification of the chemical nature and localization of the modified nucleotides even in highly abundant RNAs represents a laborious and time-consuming task. Moreover, the analysis of low abundant cellular RNAs is extremely difficult due to limited access to highly purified RNA species required for most types of analysis, like HPLC or mass spectrometry [4–7]. One alternative to this consists in direct analysis of underrepresented RNA species in total cellular RNA by reverse transcription (RT) using specific DNA primers [8]. This generally allows the sequencing of a given RNA, but the information on its modified nucleotide content is still missing. The use of specific chemical reagents reviewed in this survey explores the particular reactivity of a given modified residue and may considerably help in the interpretation of an RT profile. Another area for the use of specific chemical reactions is RNA analysis and sequencing by various types of mass spectrometry (MS), where modified residues undergo particular fragmentation pathways and their derivatization helps in identification.

2. General Reactivity of Unmodified Nucleotides

Detection of nucleotide modifications is based on either of three basic principles. For two of these, that is, separation according to physicochemical properties and differential enzymatic turnover, we refer to other reviews [4, 7, 9–12] and will mention such methods only when they are combined with the use of chemicals that specifically react with modified nucleotides. Here, we focus on the third principle, which is differential chemical reactivity. To set the

Table 1: Known RNA modifications and their abbreviations and symbols (modified nucleosides mentioned in the text are indicated in bold).

Number	Symbol	Common name
(1)	**m^1A**	**1-methyladenosine**
(2)	m^2A	2-methyladenosine
(3)	m^6A	*N*^{6}-methyladenosine
(4)	**Am**	**2′-*O*-methyladenosine**
(5)	ms^2m^6A	2-methylthio-*N*^{6}-methyladenosine
(6)	i^6A	*N*^{6}-isopentenyladenosine
(7)	**ms^2i^6A**	**2-methylthio-*N*^{6}-isopentenyladenosine**
(8)	io^6A	*N*^{6}-(*cis*-hydroxyisopentenyl)adenosine
(9)	ms^2io^6A	2-methylthio-*N*^{6}-(*cis*-hydroxyisopentenyl) adenosine
(10)	**g^6A**	***N*^{6}-glycinylcarbamoyladenosine**
(11)	**t^6A**	***N*^{6}-threonylcarbamoyladenosine**
(12)	**ms^2t^6A**	**2-methylthio-*N*^{6}-threonyl carbamoyladenosine**
(13)	**m^6t^6A**	***N*^{6}-methyl-*N*^{6}-threonylcarbamoyladenosine**
(14)	**hn^6A**	***N*^{6}-hydroxynorvalylcarbamoyladenosine**
(15)	**ms^2hn^6A**	**2-methylthio-*N*^{6}-hydroxynorvalyl carbamoyladenosine**
(16)	Ar(p)	2′-*O*-ribosyladenosine (phosphate)
(17)	**I**	**inosine**
(18)	m^1I	1-methylinosine
(19)	m^1Im	1,2′-*O*-dimethylinosine
(20)	**m^3C**	**3-methylcytidine**
(21)	**m^5C**	**5-methylcytidine**
(22)	**Cm**	**2′-*O*-methylcytidine**
(23)	**s^2C**	**2-thiocytidine**
(24)	**ac^4C**	***N*^{4}-acetylcytidine**
(25)	f^5C	5-formylcytidine
(26)	m^5Cm	5,2′-*O*-dimethylcytidine
(27)	ac^4Cm	***N*^{4}**-acetyl-2′-*O*-methylcytidine
(28)	**k^2C**	**lysidine**
(29)	m^1G	1-methylguanosine
(30)	m^2G	*N*^{2}-methylguanosine
(31)	**m^7G**	**7-methylguanosine**
(32)	**Gm**	**2′-*O*-methylguanosine**
(33)	m^2_2G	*N*^{2},*N*^{2}-dimethylguanosine
(34)	m^2Gm	*N*^{2},2′-*O*-dimethylguanosine
(35)	m^2_2Gm	*N*^{2},*N*^{2},2′-*O*-trimethylguanosine
(36)	Gr(p)	2′-*O*-ribosylguanosine (phosphate)
(37)	**yW**	**wybutosine**
(38)	o_2yW	peroxywybutosine
(39)	OHyW	hydroxywybutosine
(40)	OHyW*	undermodified hydroxywybutosine
(41)	imG	wyosine
(42)	mimG	methylwyosine
(43)	**Q**	**queuosine**
(44)	oQ	epoxyqueuosine
(45)	galQ	galactosyl-queuosine
(46)	manQ	mannosyl-queuosine
(47)	$preQ_0$	7-cyano-7-deazaguanosine
(48)	**$preQ_1$**	**7-aminomethyl-7-deazaguanosine**

Table 1: Continued.

Number	Symbol	Common name
(49)	G^+	archaeosine
(50)	Ψ	**pseudouridine**
(51)	**D**	**dihydrouridine**
(52)	**m^5U**	**5-methyluridine**
(53)	**Um**	**2′-*O*-methyluridine**
(54)	m^5Um	5,2′-*O*-dimethyluridine
(55)	$m^1Ψ$	1-methylpseudouridine
(56)	Ψm	2′-*O*-methylpseudouridine
(57)	**s^2U**	**2-thiouridine**
(58)	**s^4U**	**4-thiouridine**
(59)	m^5s^2U	5-methyl-2-thiouridine
(60)	s^2Um	2-thio-2′-*O*-methyluridine
(61)	**acp^3U**	**3-(3-amino-3-carboxypropyl)uridine**
(62)	ho^5U	5-hydroxyuridine
(63)	mo^5U	5-methoxyuridine
(64)	cmo^5U	uridine 5-oxyacetic acid
(65)	$mcmo^5U$	uridine 5-oxyacetic acid methyl ester
(66)	**chm^5U**	**5-(carboxyhydroxymethyl)uridine**
(67)	$mchm^5U$	5-(carboxyhydroxymethyl)uridine methyl ester
(68)	mcm^5U	5-methoxycarbonylmethyluridine
(69)	mcm^5Um	5-methoxycarbonylmethyl-2′-*O*-methyluridine
(70)	mcm^5s^2U	5-methoxycarbonylmethyl-2-thiouridine
(71)	nm^5s^2U	5-aminomethyl-2-thiouridine
(72)	mnm^5U	5-methylaminomethyluridine
(73)	mnm^5s^2U	5-methylaminomethyl-2-thiouridine
(74)	mnm^5se^2U	5-methylaminomethyl-2-selenouridine
(75)	ncm^5U	5-carbamoylmethyluridine
(76)	ncm^5Um	5-carbamoylmethyl-2′-*O*-methyluridine
(77)	**$cmnm^5U$**	**5-carboxymethylaminomethyluridine**
(78)	**$cmnm^5Um$**	**5-carboxymethylaminomethyl-2′-*O*-methyluridine**
(79)	**$cmnm^5s^2U$**	**5-carboxymethylaminomethyl-2-thiouridine**
(80)	m_2^6A	N^6,N^6-dimethyladenosine
(81)	Im	2′-*O*-methylinosine
(82)	m^4C	N^4-methylcytidine
(83)	m^4Cm	N^4,2′-*O*-dimethylcytidine
(84)	hm^5C	5-hydroxymethylcytidine
(85)	m^3U	3-methyluridine
(86)	**$m^1acp^3Ψ$**	**1-methyl-3-(3-amino-3-carboxypropyl) pseudouridine**
(87)	cm^5U	5-carboxymethyluridine
(88)	m^6Am	N^6,2′-*O*-dimethyladenosine
(89)	m_2^6Am	N^6,N^6,2′-*O*-trimethyladenosine
(90)	$m^{2,7}G$	N^2,7-dimethylguanosine
(91)	$m^{2,2,7}G$	N^2,N^2,7-trimethylguanosine
(92)	m^3Um	3,2′-*O*-dimethyluridine
(93)	m^5D	5-methyldihydrouridine
(94)	$m^3Ψ$	3-methylpseudouridine
(95)	f^5Cm	5-formyl-2′-*O*-methylcytidine
(96)	m^1Gm	1,2′-*O*-dimethylguanosine
(97)	m^1Am	1,2′-*O*-dimethyladenosine
(98)	**$τm^5U$**	**5-taurinomethyluridine**
(99)	**$τm^5s^2U$**	**5-taurinomethyl-2-thiouridine**
(100)	imG-14	4-demethylwyosine

Table 1: Continued.

Number	Symbol	Common name
(101)	imG2	isowyosine
(102)	ac^6A	N^6-acetyladenosine
(103)	inm^5U	5-(isopentenylaminomethyl)uridine
(104)	inm^5s^2U	5-(isopentenylaminomethyl)-2-thiouridine
(105)	inm^5Um	5-(isopentenylaminomethyl)-2′-*O*-methyluridine
(106)	$m^{2,7}Gm$	N^2,7,2′-*O*-trimethylguanosine
(107)	m_2^4Cm	N^4, N^4,2′-*O*-trimethylcytidine
(108)	m^8A	8-methyladenosine

Numbering and abbreviations of modified nucleosides are from "The RNA modification Database," http://s59.cas.albany.edu/RNAmods/.

stage, we will outline the known reactivity of the four major ribonucleotides in this paragraph. In principle, any reagent that reacts with nucleotides may be considered for chemical recognition of nucleotide modification, provided that conditions can be determined, where its reactivity significantly differs between a given major nucleotide and its modified counterpart. In the best case, conditions would be optimized to the point of exclusive reaction with either the standard or the modification. From a practical point of view, it is helpful to distinguish reagents within a narrowly defined window of experimental conditions from those which reliably and completely discriminate nucleotides over a wide plateau of experimental conditions. The former methods are evidently more difficult to newly establish in a laboratory, despite the emergence of numerous publications covering detailed protocols, and should by tackled only by experienced RNA scientists. However, even for the less complicated reactions, one should be conscious that deviations in temperature, pH, salt, incubation time, or reagent concentration may result in "leaving" the plateau and thus result in suboptimal discrimination. Similar considerations apply to the detection of the chemical species resulting for treatment with the discriminating reagent. Most often, the ease and turnaround time of detection will determine if a given reagent finds widespread use in the RNA community.

The above considerations apply to reagents of various types alike, which we have grouped into electrophiles, nucleophiles, and oxidizing, and reducing agents.

2.1. Electrophilic Reagents. Nitrogen atoms in purine and pyrimidine rings of nucleobases show varying degrees of nucleophilicity and thus react differentially with various electrophilic compounds, which are typically alkylating or acylating agents (Figure 1).

These reactivities have been explored decades ago and form the basis of chemical sequencing of end-labeled nucleic acids according to Maxam and Gilbert [13–15]. Methylating agents, in particular dimethylsulfate (DMS), preferentially alkylate the *N1* of adenosines, the *N7* of guanosines, and the *N3* of cytidines. All three sites are also found to be enzymatically methylated *in vivo*, and the methods used for sequencing of chemical probing of RNA structure can also be applied to reveal these naturally occurring modifications. An important property of many modified nucleotides is that they render oligonucleotides susceptible to chain scission via β-elimination, provided that the modification itself or an auxiliary chemical treatment ablates aromaticity of the base or leads to abasic sites. β-elimination is induced by heating of RNA with aniline at low pH. Additional treatment of m^3C with hydrazine and of m^7G [16], dihydrouridine, and wybutosine with sodium borohydride ($NaBH_4$) leads to abasic sites in RNA which are susceptible to aniline cleavage. The wybutosine base is also acid-labile and will undergo depurination upon treatment with hydrochloric acid to leave an abasic site [17]. To date, there is no satisfactory protocol to chemically reveal the presence of *N1*-methyladenosine in RNA, which therefore has to be detected by RT-based methods, as will be detailed below.

Other popular electrophilic agents that react with nitrogens include diethyl pyrocarbonate (DEPC, *N7* of adenosine) and ethylnitrosourea, which alkylates all nitrogens in addition to phosphates and all oxygens in nucleic acids [18]. *N*,*N*-(dimethylamino) dimethylchlorosilane (DMAS-Cl) reacts with the N^2 of guanosines [19]. Certain reagents gain in specificity due to the presence of two electrophilic groups, for example, chloroacetaldehyde (*N1*-A, N^6-A *N3*-C, N^4-C) [20, 21], chlorotetrolic (4-chloro-2-butynoic) acid ester (*N1*-A, N^6-A *N3*-C, N^4-C) [22], and glyoxal and kethoxal (*N1*-G, N^2-G) [23, 24]. Finally, carbodiimides, in particular 1-cyclohexyl-3-(2-(4-morpholinyl)ethyl) carbodiimide tosylate (CMCT), acylate the *N3* of uridines and the *N1* and N^2 of guanosines, as well as both nitrogens in pseudouridine and *N3* in inosine. Its use for the mapping of Ψ residues will be detailed below.

2.2. Oxidizing Agents. As electron poor species, oxidizing agents share certain properties and also target moieties with alkylating agents. Nitrosyl cations and related species react with exocyclic nitrogens in cytidine, adenosine and guanosine to form diazonium compounds, which eventually hydrolyze and yield the corresponding deamination products: uridine, inosine, and xanthene, respectively. Monoperphtalic acid was reported to oxidize adenosine to yield adenosine 1-*N*-oxide [25]. The Δ5,6 double bonds in pyrimidines can in principle be oxidized by hydrogen peroxide although permanganate and in particular osmium-VIII compounds are most frequently used (reviewed in [26]).

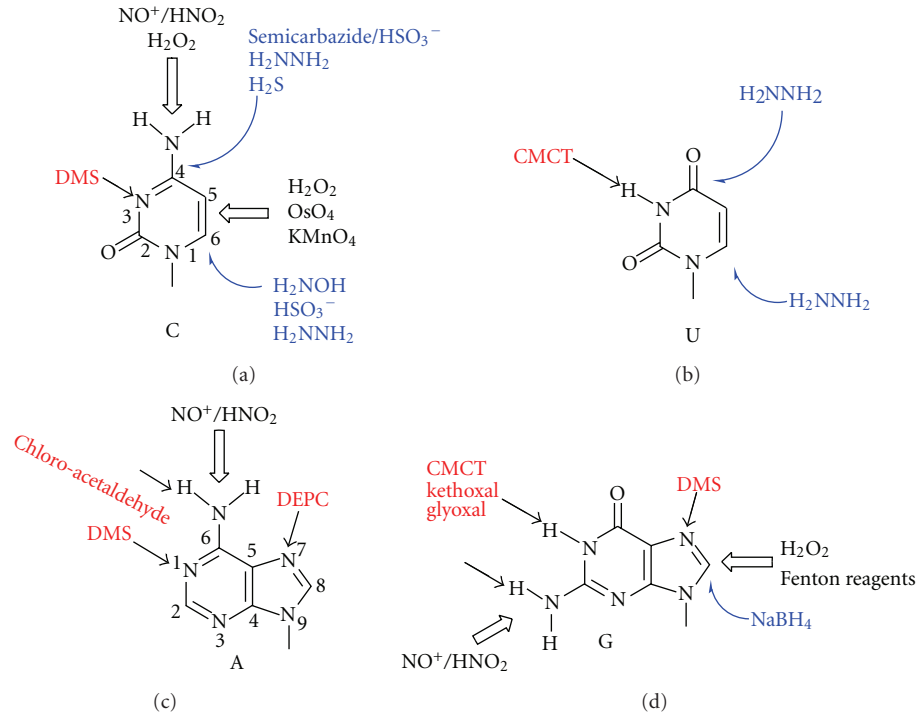


Figure 1: General chemical reactivity of nucleobases. Attack sites on the nucleobase are indicated by arrows. Open black arrows indicate oxidizing agents, filled black ones indicate alkylating electrophiles, and blue ones indicate nucleophiles. Numbering of the nucleobase atoms 1–6 for pyrimidines or 1–9 for purines is indicated outside the ring.

Radical generating species, including Fenton reagents [27], copper phenantroline [28, 29] and peroxonitrous acid [30] are known to abstract hydrogens from CH bonds in the ribose. So far, this is the only type of reactivity that has not been exploited for the detection of nucleotide modifications. This type of reagents also causes the formation of 8-oxo-guanosine, which is the major lesion in oxidative damage of DNA [31].

2.3. Nucleophilic and Reducing Reagents. In pyrimidines, the Δ5,6 double bond is part of a Michael acceptor, which can be attacked in the 6 position by strong nucleophiles such as, bisulfite, hydrazine, methoxamine, hydroxylamine [32–34], and bisulfite [35–37] but also by KI/$TiCl_3$ [38]. $NaBH_4$ will act only on methylated *N3*-C, as pointed out already.

The C4 position is also electrophilic, but typically less reactive than the Michael acceptor electrophile C6. Nucleophilic attack here is known from the bidentate hydrazine [39], or a combination of bisulfate and a semicarbazide [40]. In both cases, the first nucleophilic attack occurs at the C6, and only then does the C4 position react [33, 39]. Using hydrogen sulfide under high pressure, cytidines can be converted into s^4U as the result of a nucleophilic attack at the carbon 4 position [41, 42], which might be preceded by an initial nucleophilic attack at the C6.

3. Specific Reactivity of Modified Nucleotides

From a perspective of chemical reactivity, it is helpful to distinguish between modifications carrying principally new chemical moieties, and those which only moderately alter the reactivity of chemical moieties already present in RNA. The latter category essentially contains all methyl group additions, and reagents and reaction conditions must be carefully controlled to achieve discrimination. Thus, the detection of m^3C, which has already been mentioned (see Figure 2(a)), is based on the susceptibility of cytidines towards nucleophiles, which was increased by the methyl group modification to *N3*, and thus allowed for detection after hydrazine treatment. In contrast, addition to the 5 position of a methy lgroup renders cytidines more electron rich and consequently less electrophilic. Therefore, nucleophilic attacks by bisulfite can be restricted to cytidines under carefully controlled conditions which leave m^5C unaffected by bisulfite.

A number of more sophisticated modifications involve the introduction of chemical moieties which are chemically orthogonal, that is, they can be conjugated by certain reagents to which unmodified RNA is completely inert. Thus, isothiocyanate and NHS-derivatives selectively react with primary amines as present, for example, in acp^3U, or lysidine (k^2C). Thiolated nucleotides such as s^2U, s^4U, and s^2C [43, 44] react with iodoacetamide derivatives, and carboxylic acids such as in t^6A/m^6t^6A and acp^3U [45, 46], other COOH-containing nucleotides can most probably be activated by carbodiimides for conjugation with a nitrogen nucleophile.

3.1. 5-Methylcytosine. The chemistry of m^5C in DNA has been subject to intense research as a consequence of the

importance of m^5C in the field of epigenetics. Bisulfite m^5C sequencing is a state-of-the art detection method and used on a very large scale. The addition of bisulfite to the 6 position promotes a series of chemical transformations depicted in Figure 2(b), which ultimately leads to deamination. Cytidines are thus transformed into uridines and read as such in subsequent sequencing reactions. Based on their resistance to deamination reaction, sequencing after bisulfite treatment will reveal m^5C residues as cytidine signals. Despite lasting efforts, adaptation of bisulfite to RNA has only recently succeeded [35–37] and is now one of the few methods that might be used to investigate RNA modification on a genome-wide scale [47, 48].

As indicated in Figure 2(c), another avenue of chemical discrimination between cytidines and m^5C may involve oxidation of the double bond. Although the oxidation with permanganate produces insoluble MnO_2, and further oxidation of the glycol to bis-aldehyde may occur under certain conditions, it has been successfully employed in the discrimination of deoxy-m^5C versus deoxycytidine [49, 50]. The osmium tetroxide reactions do not involve the formation of insoluble products, but the reagent is highly toxic and mutagenic. In the past years, a variety of osmate-based bioconjugate reagents have been developed for the selective detection of m^5C in DNA. These reagents exploit the additional electron density brought about by the methyl group for highly selective formation of a stable complex with deoxy-m^5C but not with deoxycytidine. However, applications to RNA are still lacking [51–53] (reviewed in [26]).

3.2. 3-Methylcytosine. Detection of m^3C is based on its increased reactivity towards hydrazine, as discussed above. Figure 2(a) shows the reaction sequence leading to an intermediate susceptible to aniline-induced cleavage *via* β-elimination. Cleavage sites can be revealed by sequencing reactions with end-labeled RNA [14].

3.3. 7-Methylguanosine. Guanosine methylated at position 7 (m^7G) is frequently present in the variable region of tRNAs. As has been pointed out, its detection can be achieved by aniline-induced cleavage of the RNA chain by β-elimination after additional treatment under alkaline conditions [16] or after its reduction by sodium borohydride ($NaBH_4$, shown in Figure 3). Because solutions of $NaBH_4$ are typically alkaline, and because the reduction product has been described as sensitive to reoxidation in air, the actual species undergoing beta elimination may not be well defined. This reaction was first studied for isolation of defined tRNA fragments [54] and later applied to RNA sequencing using DMS in Maxam and Gilbert type chemistry [14] or RNA structural probing.

3.4. 2′-O-Methylated Riboses. Alkaline hydrolysis of RNA polynucleotide chains proceeds via deprotonation of the ribose 2′-OH to the corresponding alcoholate, and its nucleophilic attack of the nearby 3′-phosphate. The resulting intermediate is unstable and rapidly decomposes into a 2′,3′-cyclophosphate leading to phosphodiester bond cleavage (Figure 4(a)). The methylation of ribose 2′-OH prevents alcoholate formation, and thus decreases the reactivity of the 2′-oxygen and consequently prevents phosphodiester bond cleavage almost completely [55, 56]. This resistance to alkaline hydrolysis of the phosphodiester bonds on the 3′-side of 2′-*O*-methylated nucleotides can be detected by primer extension or by direct analysis of the cleavage profile of end-labeled RNA. The 2′-*O*-methylated residue appears as a "gap" in the regular ladder of OH^- cleavage.

A very sensitive method involving RNase H digestion directed by 2′-*O*-methyl RNA-DNA chimeras can be used to confirm the presence of a 2′-*O*-methylated residue. RNase H nonspecifically cleaves the RNA strand of an RNA-DNA hybrid. However, it does not cleave any site where the 2′-*O* position of an RNA residue is methylated. The digestion of an RNA directed by a complementary 2′-*O*-methyl RNA-DNA chimeric oligonucleotide consisting of four deoxynucleotides flanked by 2′-*O*-methyl ribonucleotides is site specific. The use of a 2′-*O*-methyl RNA-DNA chimera and RNAse H digestion therefore provide a direct assay for determining whether any particular nucleotide in a long RNA molecule carries a 2′-*O*-methyl group. If the 2′-*O* position of the targeted residue is methylated, RNAse H cleavage is blocked. If instead the 2′-*O* position is unmodified, RNAse H should cut the RNA specifically at that site [57].

Another method based on the RNA treatment by sodium periodate followed by β-elimination has been recently used to analyze the methylation status of the 3′-terminal nucleotide of plant miRNAs [58, 59]. The reactions eliminate a 3′ nucleotide containing both 2′- and 3′-OH groups on the ribose (Figure 4(b)), resulting in an RNA product that is one nucleotide shorter than the substrate RNA and that contains a 3′ phosphate group [60]. Therefore, miRNAs that contain a 3′ terminal ribose with both 2′- and 3′-OH groups will migrate faster during electrophoresis between one and two nucleotides after the treatment. Methylated miRNAs whose 3′ terminal ribose contains a 2′-*O*-methyl group do not participate in the reactions and remain unchanged in mobility. A significant disadvantage of this approach is that it does not distinguish a 2′-*O*-methyl modification from other potential modifications on the 3′ terminal ribose. In studies by Yu et al., it was possible to assign the identity of the 3′-terminal nucleotide modification thanks to mass spectrometry studies [58].

More recently, an alternative chemical treatment which should allow the distinction between unmethylated and 2′-*O*-methylated residues has been proposed [61, 62]. Free unconstrained ribose 2′-OH easily reacts with N-methylisatoic anhydride (NMIA, Figure 4(c)), and the resulting adduct can be analyzed by primer extension (selective 2′-hydroxyl acylation analyzed by primer extension, SHAPE). For the moment, the method was only applied to RNA structural analysis. However, comparisons of NMIA modification and primer extension profiles for unmodified transcript and 2′-*O*-methyl-containing RNA may probably be used as a tool for 2′-*O*-methylation mapping.

Yet another approach allowing the separation of O-methyl nucleosides from ribose-unsubstituted nucleosides in one chromatographic step has been developed in the early 80s [63]. This methods is based on the capability of boronate

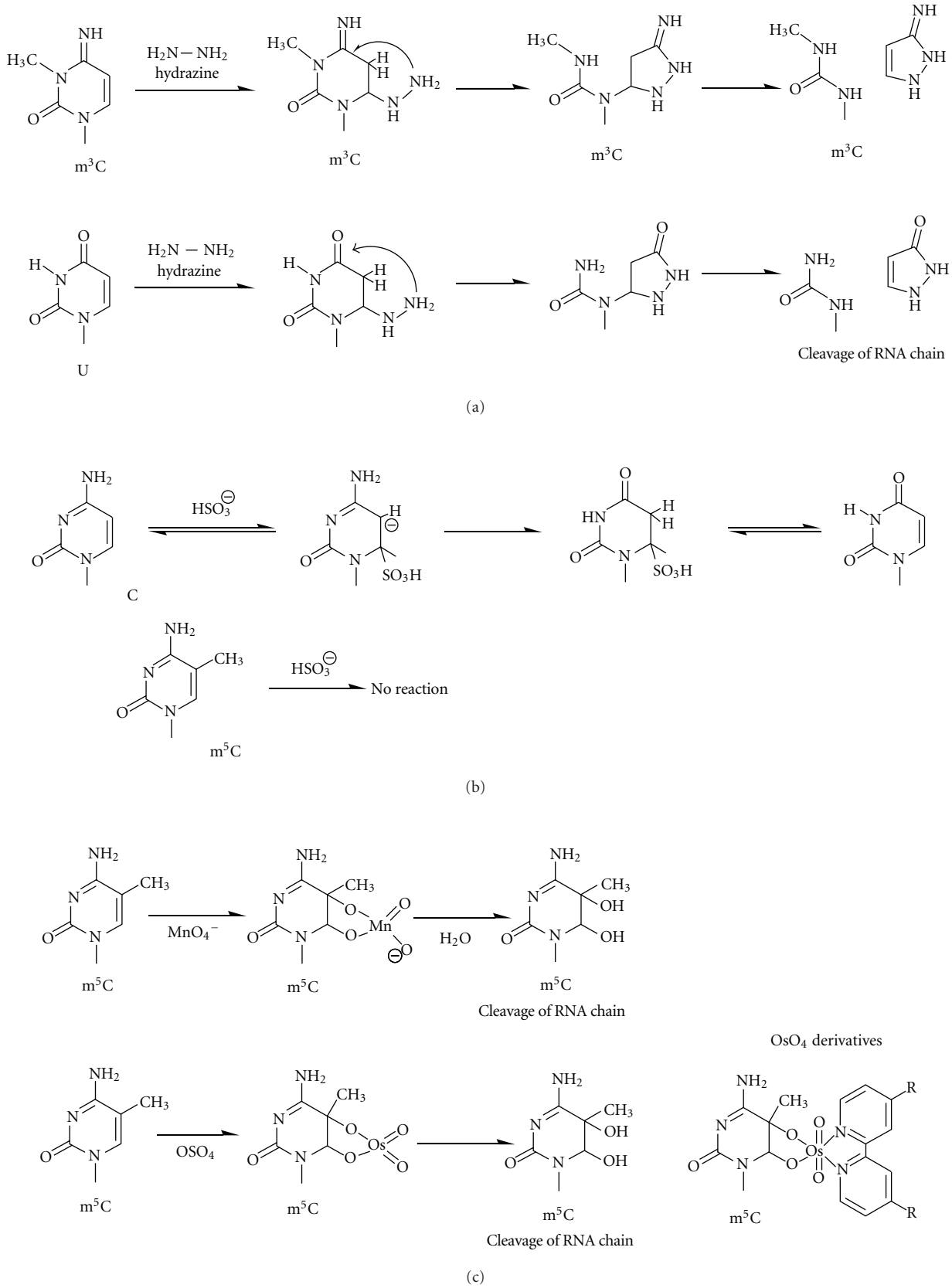


Figure 2: Specific chemical reactions for m^3C and m^5C detection. (a) Cleavage of m^3C and unmodified uridine by hydrazine. (b) Deamination of unmodified cytosine by bisulfate; m^5C is resistant for deamination. (c) Specific oxydation of 5-6 double C=C bond in methylated pyrimidines by MnO_4^-, OsO_4, and OsO_4 derivatives.

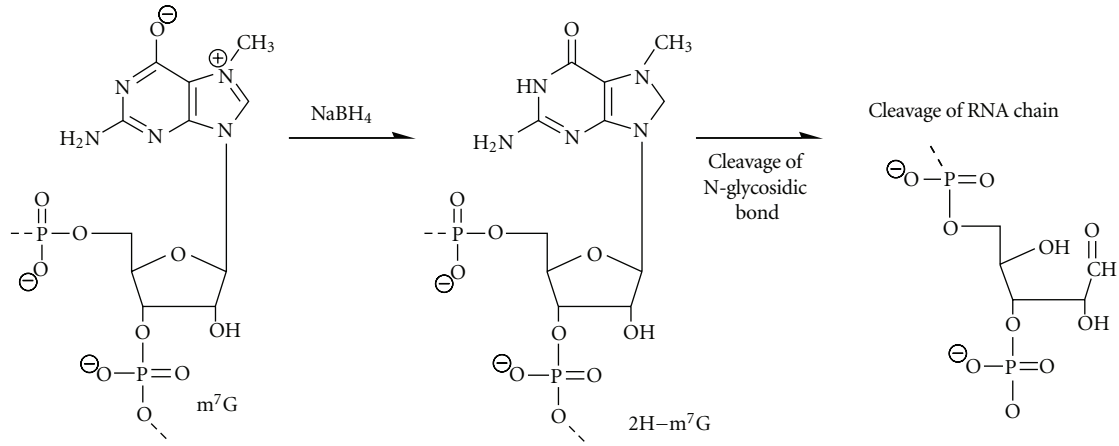


Figure 3: Sodium borohydride ($NaBH_4$) reduction of m^7G in RNA. Reduction of m^7G leads to formation of abasic site in RNA followed by the cleavage of the RNA chain by β-elimination.

to form an anionic complex with the cis-2',3' hydroxyls of unsubstituted ribonucleosides, but not with ribose-substituted nucleosides such as 2′-*O*-methylnucleosides. Phenyl boronates with hydrophobic side chains of about 1 nm in length have been synthesized and used to coat inert 10 μm beads of polychlorotrifluoroethylene. This matrix complexes easily with compounds containing free vicinal cis-hydroxyls allowing their separation from their O-alkyl or O-acyl derivatives. Boronates also find application in affinity electrophoresis [64], as will be detailed later on.

3.5. Pseudouridine. Pseudouridine (psi, Ψ) is the result of an isomerization reaction during which the *C5* and N1 positions of uracil are interconverted. The resulting nucleotide features a glycosidic carbon-carbon bond and an additional amide functionality, which is available for chemical discrimination (Figure 5). Because Ψ retains the base pairing characteristics of uridine, it is invisible in RT profiles and must be chemically modified (Figure 5(a)) for detection by either mass spectroscopy [65–67] or by RT [8]. Somewhat surprisingly, chemical modification can be achieved with satisfying selectivity by two classes of electrophiles. Michael acceptors such as acrylonitrile [68] or methylvinylsulfone [69] alkylate the *N3* while not reacting with the four major nucleotides (Figure 5(b)). These reagents are, however, known to react also with inosine, m^5C, and m^5U [69, 70].

Ψ residues can be acylated on both nitrogens by carbodiimides (Figure 5(a)) [71, 72]. The reaction also acylates both the *N3* of uridines and the *N1* of guanosines, but all residues except for the one on *N1* of Ψ can be removed by subsequent alkaline treatment [73, 74]. Indeed, whereas U and G adducts of CMC are readily cleaved by weakly alkaline conditions, cleavage of N3-CMC-Ψ requires 7 M NH_4OH at 100°C for 8 min [73]. The remaining bulky CMC residues are then detected by primer extension using reverse transcriptase.

Recent extensive studies on pseudouridine mapping by CMCT treatment and MALDI-TOF MS pointed out that modified nucleotide ms^2i^6A present in *E. coli* tRNA also retains a CMCT derivative after acylation [67]. However, the chemical basis for this particular reactivity was not elucidated in detail.

Another approach used to detect Ψ residues depends upon the greater resistance of Ψ and m^5U residues to hydrazinolysis compared to U and most other U-derived bases (Figure 5(c)). The reaction with hydrazine leads to pyrimidine ring opening (see Figure 2(a)). Subsequent aniline treatment cleaves the polynucleotide chain leading to the termination of RT [75–77]. Ψ residues, being resistant to hydrazinolysis, are not cleaved and do not stop RT.

3.6. Thiolated Nucleotides. Sulfur atoms, even when introduced into RNA as thiocarbonyl functions, are easily oxidized but are also very nucleophilic and thus react with a variety of electrophiles under conditions under which the remainder of RNA residues stays inert. Iodoacetamide derivatives, for example, of fluorescent dyes, are typical reagents to derivatize sulfur residues in proteins, that is, free cysteins. Diluted solutions of these reagents have been shown to selectively conjugate to thiolated nucleotides such as s^2U, s^4U, and s^2C [43, 44] (Figure 6). Thiouridines have also been reported to react with carbodiimides like CMCT [78], and the resulting conjugates can be identified by mass spectroscopy [65, 67].

The thiophilic character of mercury ions has been exploited for affinity electrophoresis in polyacrylamide gels copolymerized with (*N*-acryloylamino)phenyl-4-mercuric chloride (APM-gels, Figure 6). Sulfur-containing RNAs are retarded during migration, and interestingly, the degree of retardation depends on the nature of the thiomodification; s^4U is more strongly retarded comparing to s^2U [79].

3.7. Modifications Containing Primary Amines. Primary amines, which are not present in unmodified RNA, are relatively strong nucleophiles which can be selectively conjugated to a number of commonplace amino-derivatizing agents

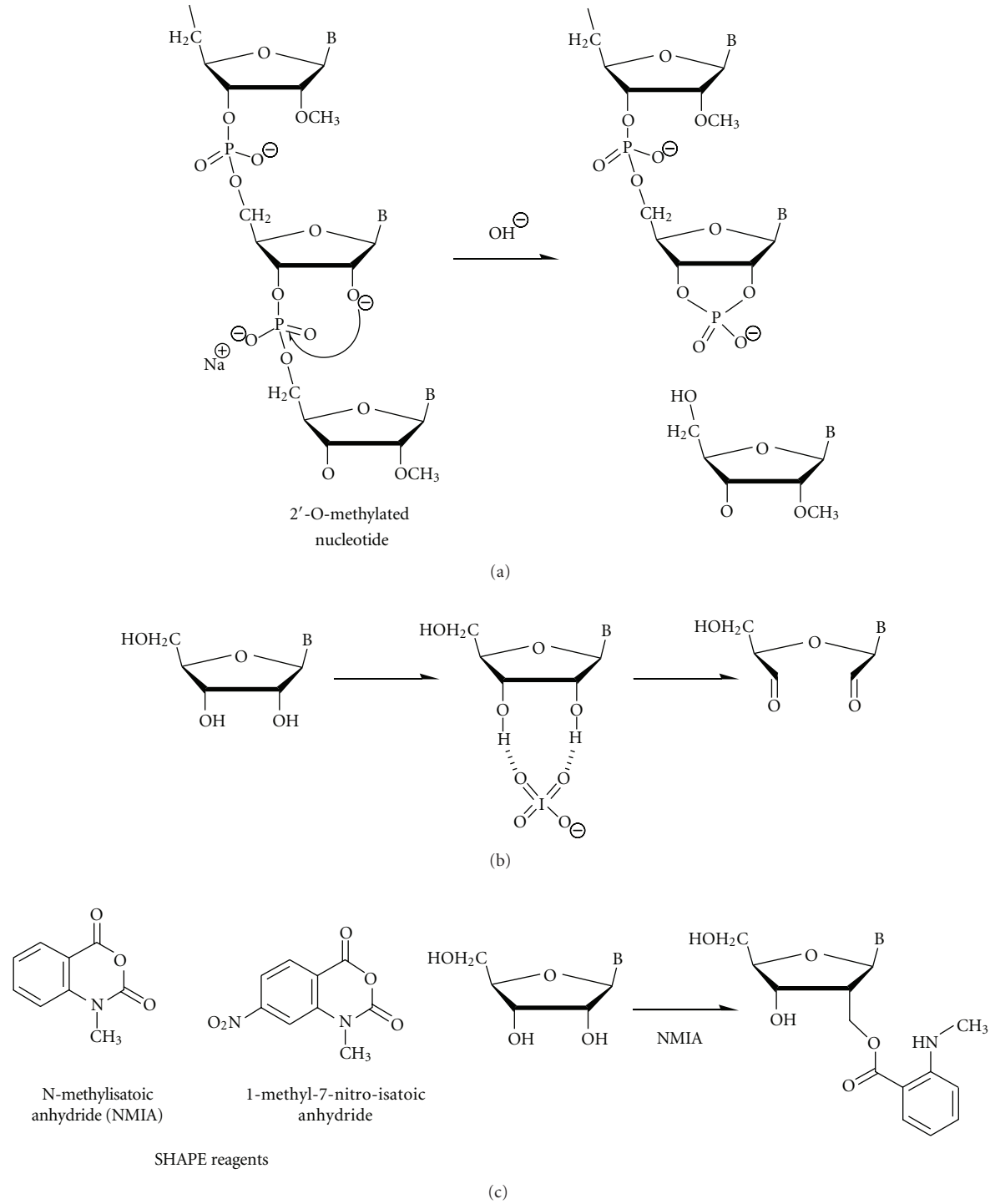


Figure 4: Specific reagents for 2′-*O*-methylated residues in RNA. (a) Selective cleavage of nonmethylated ribose residues at alkaline pH. The 2′-*O*-methylated residues are resistant for such cleavage. (b) Periodate oxydation of cisdiol residues at the terminal 3′-ribose. (c) Acylation of ribose free 2′-OH by SHAPE reagents N-methylisatoic anhydride (NMIA) and 1-methyl-7-nitro-isatoic anhydride.

containing, for example, *N*-hydroxysuccinimide and isothiocyanate moieties [78, 80, 81], and probably with pentafluorophenol derivatives, too (Figure 7(a)). For example, labeling with fluorescent dyes has been successfully implemented for *E. coli* $tRNA^{Phe}$ containing acp^3U [78, 81]. Fluoresceinisothiocyanate under mild alkaline conditions was successfully applied for the labeling of *E. coli* $tRNA^{Tyr}$ at the queosine position [82]. Free amino functions may also react with

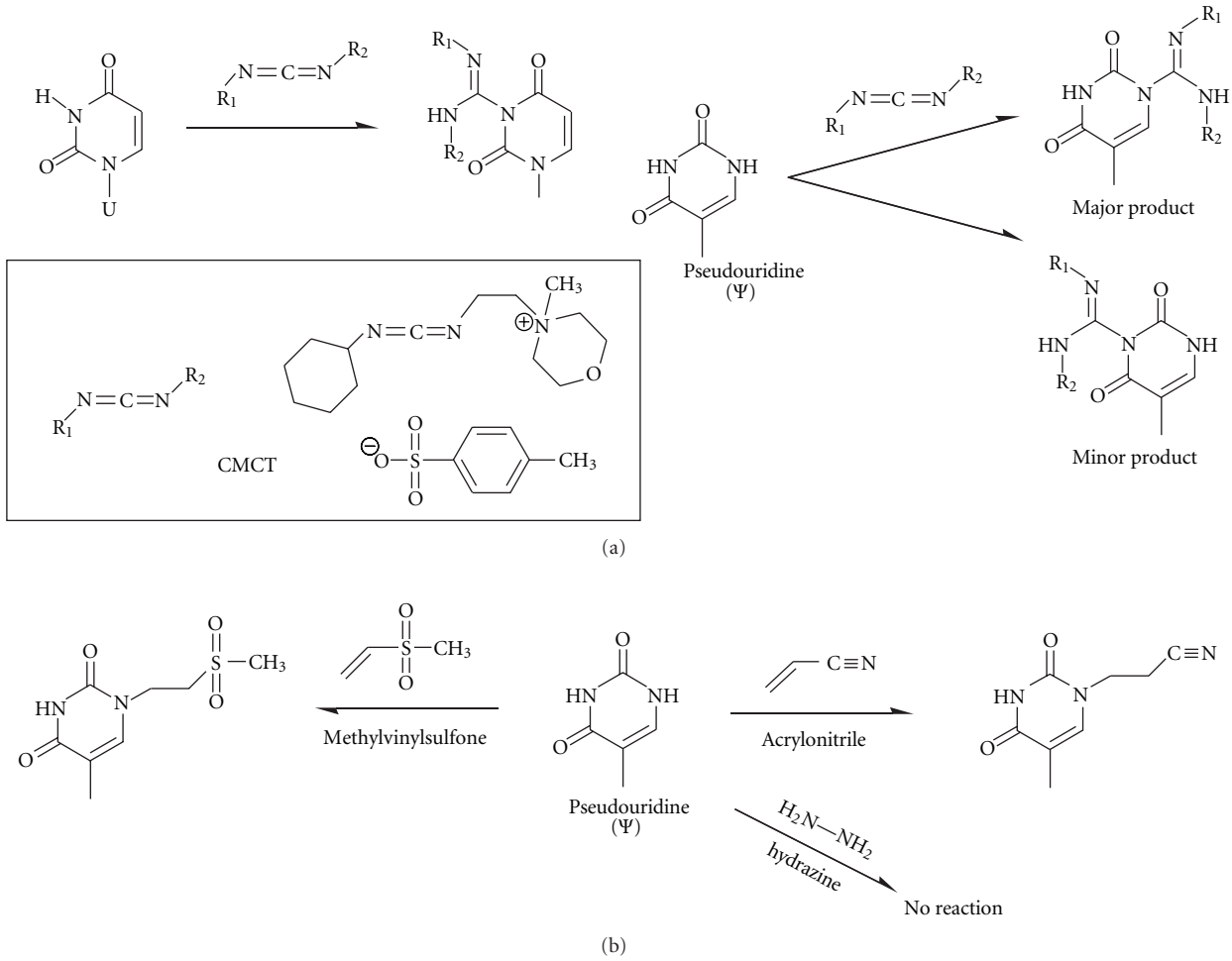


Figure 5: Pseudouridine (Ψ-) specific reagents. (a) Pseudouridine and uridine reactivity with CMCT and similar water-soluble carbodiimides. Structural formula of CMCT is shown on the left. (b) Pseudouridine reactions with methylvinylsulfone and acrylonitrile. Pseudouridine residues are resistant to hydrazine, while unmodified uridines are specifically cleaved (see Figure 2(a)).

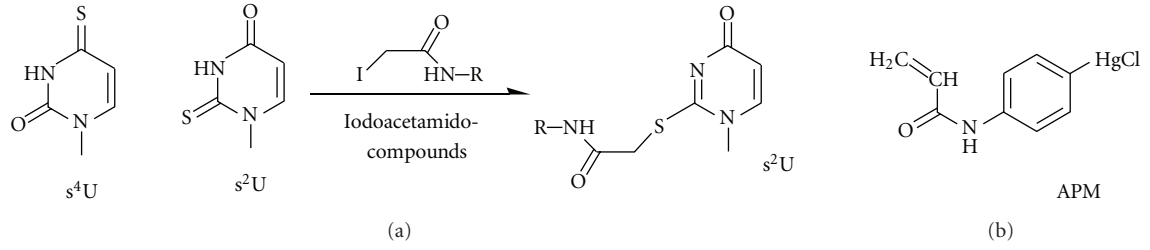


Figure 6: Examples of thiolated nucleotides found in RNA (a) and (*N*-acryloylamino)phenyl-4-mercuric chloride (APM) which is used for gel-retardation detection of thiolated molecules. (b) Reaction of thiolated nucleotides with iodoacetamide and its derivatives. s^2U is taken to illustrate the reaction product.

dansyl chloride under neutral conditions, as was demonstrated by selective labeling of $preQ_1$-34 in the same *E. coli* $tRNA^{Tyr}$ obtained from queosine-deficient strain [83] (Figure 7(a)).

3.8. Modifications Containing free Carboxylic Acid Groups. A number of modifications typically occurring at positions 20, 34, and 37 of tRNAs consist in the incorporation of an amino acid via its amino function, resulting in the presence in the modified RNA of a free carboxy group. This can be selectively activated for conjugation with amine reagents by dilute water-soluble carbodiimide, which will not react with the other nucleobases under these conditions (Figure 7(b)). The principle has been demonstrated with

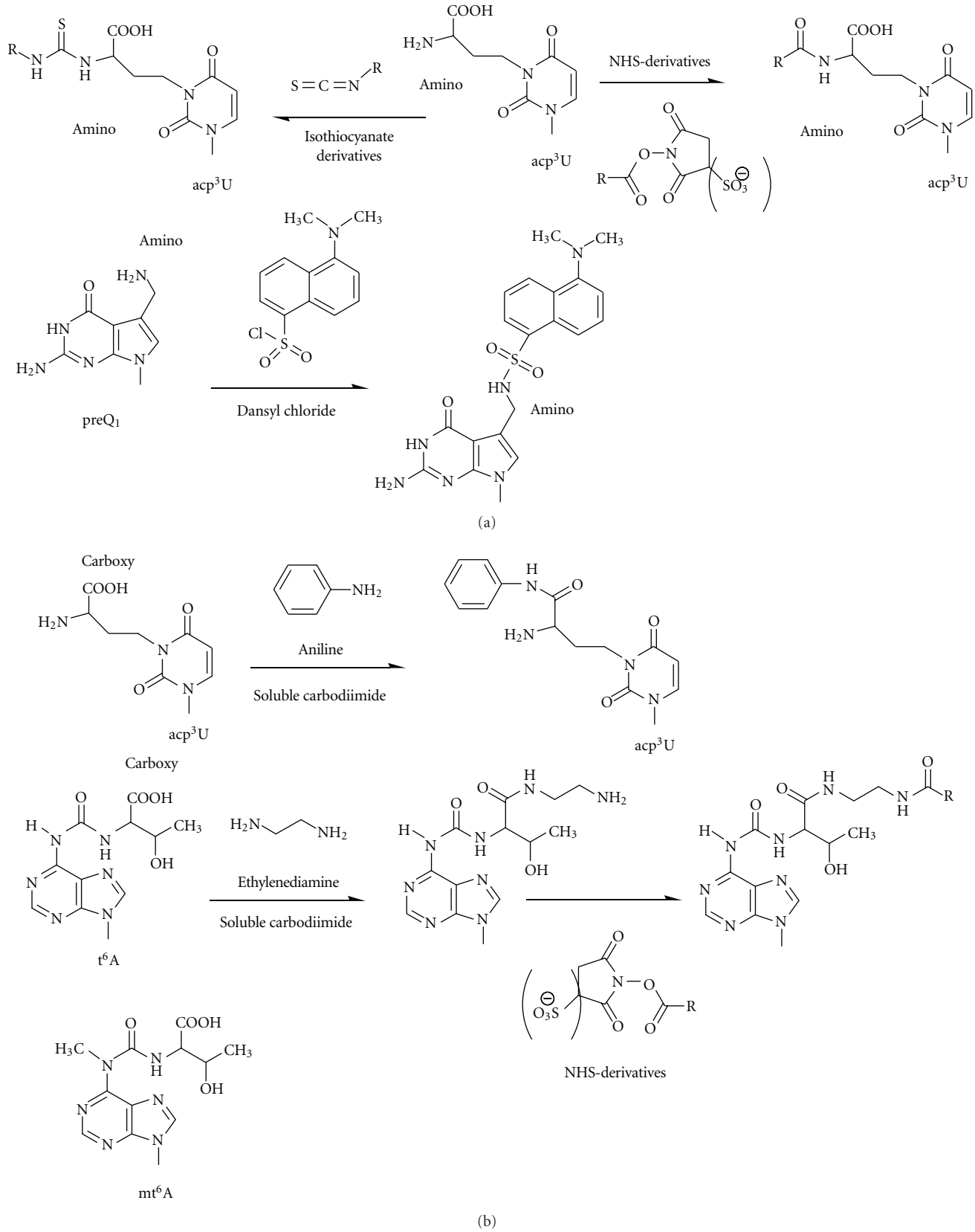


FIGURE 7: Specific reactivity of free NH_2– (a) and –COOH (b) groups in modified nucleotides. acp^3U is taken to illustrate the reactivity with isothiocyanate derivatives and NHS derivatives. Reaction with dansyl chloride is shown for $preQ_1$ (a). Reactions of free –COOH groups with aniline, ethylenediamine and similar molecules in the presence of soluble carbodiimide. The resulting free NH_2-group of ethylenediamine may be further used for attachment of activated acyl (b). The structure of mt^6A(m^6t^6A) is shown at the bottom.

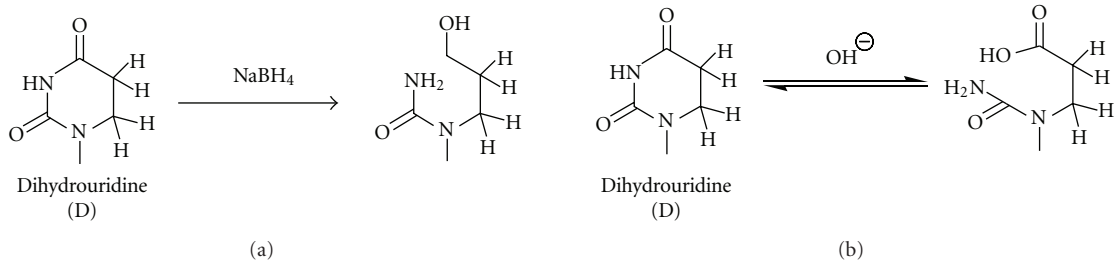


Figure 8: Cleavage of the dihydrouridine ring (a) upon reduction by sodium borohydride ($NaBH_4$) and (b) at mild alkaline conditions. In both cases, cleavage of the dihydrouridine ring is followed by cleavage of the RNA chain.

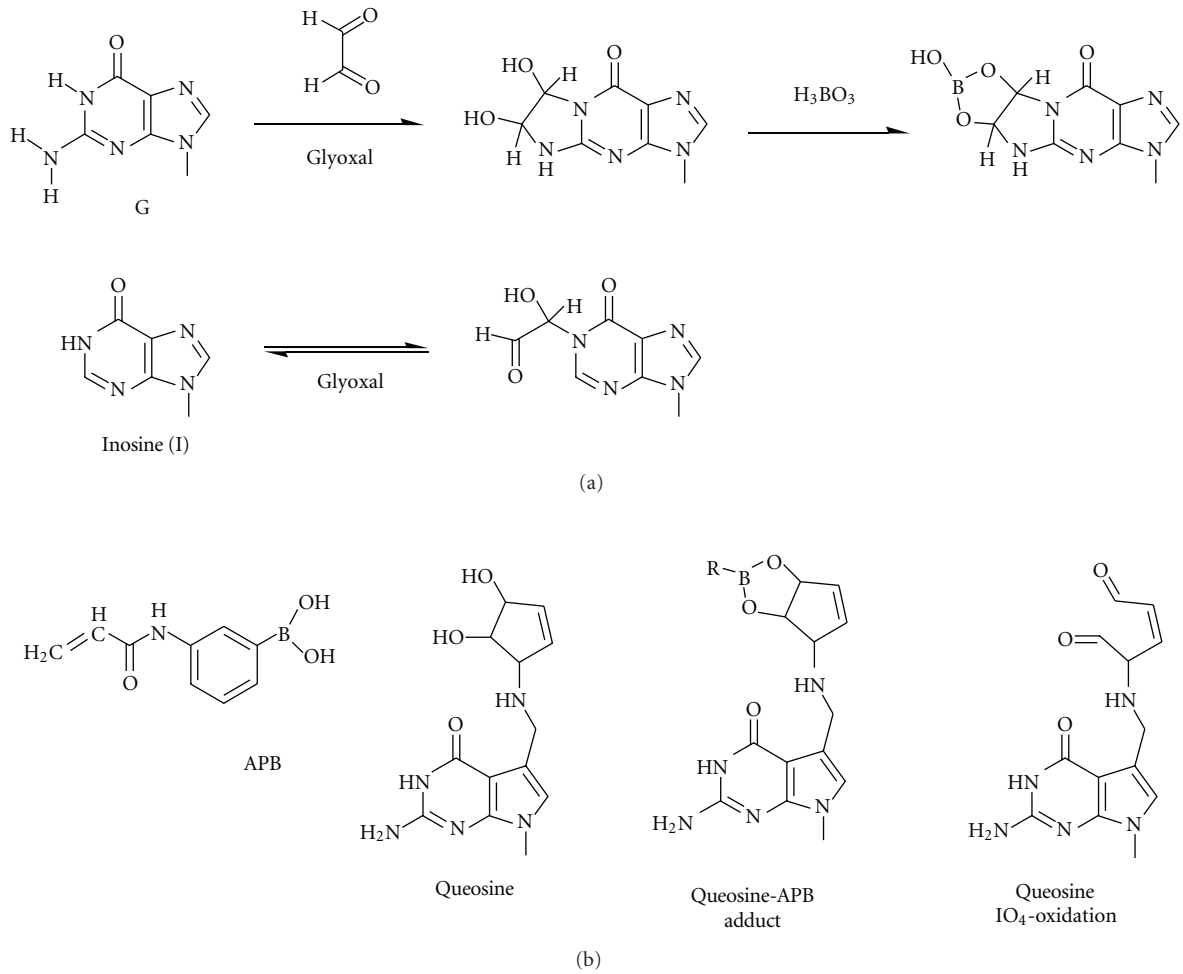


Figure 9: Reaction of guanosine residues in RNA with glyoxal and stabilization of the resulting adduct by boric acid H_3BO_3. Inosine residues do not form the stable product under similar conditions (a). Organic derivative of boric acid (*N*-acryloylamino)phenyl-3-boronic acid (APB) and its complex with cisdiols present in queosine and the dialdehyde resulting from periodate oxydation of the cisdiol (b).

mt^6A (m^6t^6A) and acp^3U [45, 46] but may be applicable to a variety of other modifications, potentially including ms^2t^6A, m^6t^6A, g^6A, hn^6A, and ms^2hn^6A, chm^5U, cmo^5U, $cmnm^5s^2U$, $cmnm^5Um$, $cmnm^5U$, $m^1acp^3\Psi$, τm^5U, and τm^5s^2U and possibly others.

3.9. Dihydrouridine. In contrast to all other pyrimidines, the dihydrouridine (D) has a saturated pyrimidine ring due to reduction of the double C=C bond between carbons 5 and 6. This considerably affects the stability of D both in the reduction by $NaBH_4$ [84–86] and under alkaline

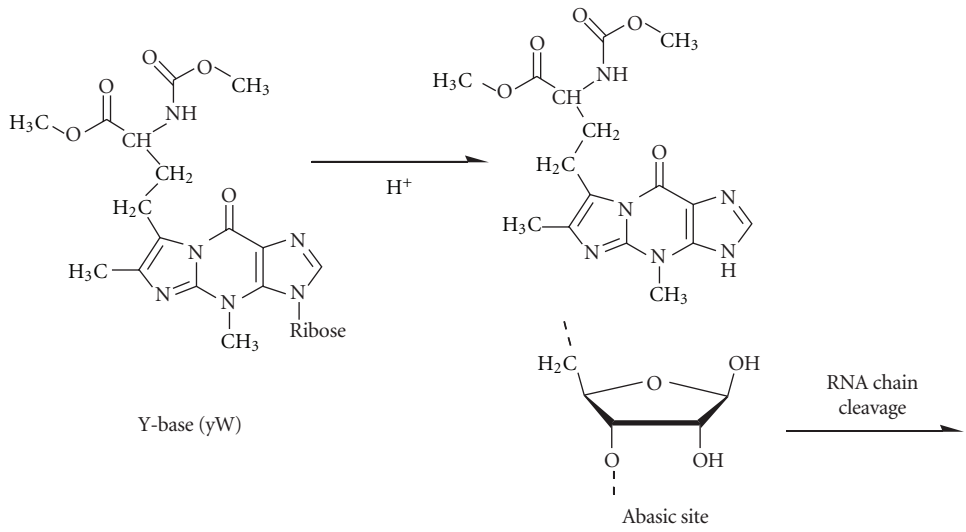


Figure 10: Depurination of wybutosine base (yW, Y-base) under mild acidic conditions. The formation of abasic site in RNA leads to the subsequent cleavage of the RNA chain.

conditions [87]. Initial studies of tRNA nucleosides by $NaBH_4$ treatment revealed that D, ac^4C, and s^4U are reduced under rather harsh reaction conditions [84], while major nucleosides, as well as Ψ residues, were found to be inert. The product of D reduction by $NaBH_4$ is a ureidopropanol riboside (Figure 8(a)). Further studies demonstrated that the D residue in tRNA can be selectively reduced by $NaBH_4$ under very mild conditions at pH 7.5 in Tris-HCl buffer on ice [86].

Another characteristic of D is its particular sensitivity to alkaline hydrolysis. Hydrolytic ring opening (Figure 8(b)) of dihydrouridine occurs quite rapidly at elevated temperatures and a pH above 8.0–8.5 [88]. This instability may explain the absence of this modification in hyperthermophiles and was also used for detection of D residues in yeast tRNA and characterization of the corresponding tRNA D-syntases [87].

3.10. Inosine. The modified nucleotide inosine (I) is derived from adenosine (A) by enzymatic deamination catalyzed by specific adenosine deaminases [89, 90]. RNase T_1 does not distinguish inosine residues present in RNA from guanosine residues and cleaves on their 3′ side. However, preliminary specific reaction with glyoxal allows RNase T_1 to distinguish G and I residues [91] (Figure 9(a)). Indeed, glyoxal reacts with *N1* and N^2 of G residues and thus abolishes their recognition by RNase T_1. The resulting covalent adduct is stabilized by boric acid. By contrast, inosine cannot react in the same way with glyoxal due to the absence of NH_2-group at position 2 and the glyoxal-inosine adduct is unstable. Thus, RNase T_1 cleavage of glyoxal-treated RNA proceeds only at inosine residues and not at G residues. The cleavage positions may be detected by primer extension analysis or other methods.

3.11. Queosine. The complicated hypermodification queuosine contains two vicinal hydroxyl groups, a feature otherwise only present in the most 3′-ribose of an RNA chain. A chelating effect of these hydroxyl groups can be exploited by complexation to Lewis acids such as boronic acid. Igloi and Kössel [64] have exploited this feature to develop affinity electrophoresis based on the incorporation of (*N*-acryloylamino)phenyl-3-boronic acid (APB) into polyacrylamide gels (Figure 9(b)). Affinity chromatography based on the same principle has been described by Vogeli et al. [92]. Of note, vicinal hydroxyl groups can be selectively oxidatized to yield dialdehydes which can be selectively condensed with hydrazine derivatives, yielding hydrazones that can be further stabilized by reduction with borohydrides.

3.12. Wybutosine. The reaction of the wybutosine base (yW) in yeast $tRNA^{Phe}$ with $NaBH_4$ has already been mentioned. Its depurination under mild HCl treatment (pH 2.9, 37°C, 2-3 hours, Figure 10) was already noticed in the late 60s and 70s [17, 93]. Later on, the reaction of the riboaldehyde group formed at the resulting abasic site with hydrazine and hydrazine derivatives was used for specific fluorescence labeling of $tRNA^{Phe}$ for biophysical studies of the ribosomal translation [86, 94].

4. Practical Use of Specific Reagents for Detection of Modified Nucleotides by RT or RNA Chain Cleavage

Specific chemicals like those described above may be used both for detection and the precise localization of modified residues in RNAs. The techniques developed for RNA sequencing using Maxam and Gilbert's approach also allow for detection of modified nucleotides, since many chemical reactions that are specific for modified nucleotides may be exploited to induce cleavage of RNA polynucleotide chain, often in combination with further chemical treatment. This is the case, for instance, for hydrazine cleavage, $NaBH_4$ reduction, or hydrolysis of phosphodiester bonds under

various conditions. Analysis of the cleavage position can be performed with 5′-^{32}P-labeled RNA and separation of the cleaved fragments by electrophoresis on denaturing PAGE. However, these techniques require difficult and time-consuming purification to homogeneity of a sufficient amount (at least 0.5–1 μg) of a given RNA species and thus cannot be applied to low-abundance cellular RNAs.

An alternative approach for the detection of RNA modifications is based on the use of RT. A synthetic DNA oligonucleotide designed to target a specific sequence is annealed to cellular RNA without previous RNA purification. Primer elongation by reverse transcriptase allows both sequencing of RNA species by Sanger type reactions and, in addition, the detection of specific reagent-dependent stops, which are indicative of the presence and exact position of modified nucleotides. A limited number of natural RNA modifications altering Watson-Crick base pairing can be detected by this approach even in the absence of a preceding chemical treatment. For RT-silent modifications, a specific reagent is required to create a bulky chemical modification, which blocks base pairing and consequently progression of the reverse transcriptase, thus resulting in an RT stop. Alternatively, cleavages of the RNA chain generated by specific treatment may also be detected by primer extension. However, interpretation of RT profiles should be undertaken with caution, since the presence of stops or gaps in the RT profile is frequently hidden by natural pauses of the RT at the corresponding positions.

For some RNA modifications, a simple primer extension under specific conditions may also be used for detection. Such method was developed for the precise mapping of 2′-*O*-methylated residues in RNA and uses primer extension by reverse transcriptase at low dNTP concentrations [55, 56]. It is based on the observation that limited dNTP concentrations cause the reverse transcriptase to frequently slow down (or pause) at such nucleotides. Primer extension with normal (unmodified) residues is much less sensitive to reduced dNTP concentration. Many other modified nucleotides in RNA do not create such pauses, and thus low dNTP concentration-dependent stops seem to be rather specific to 2′-*O*-methylation. The comparison of RT profile for natural (modified) RNA and unmodified RNA transcript may help to discriminate modification-dependent from other RT stops. It is noteworthy that pauses observed at 2′-*O*-methylated residues are sequence dependent and, depending on the sequence context, no pause may occur at some of the 2′-*O*-methylated residues. Therefore, the method does not allow an exhaustive identification of all 2′-*O*-methylated residues in RNA molecules. In some instances, RNA sequence changes resulting from RNA editing (e.g., adenosine or cytidine deamination) can also be directly detected by comparison of RT profile for unmodified transcript and natural RNA species.

Other methods employing RT are based on particular base pairing properties of some modified nucleotides. An alternative approach for inosine detection consists in direct RNA sequencing with RT and comparison of the cDNA sequence obtained with the corresponding genomic sequence. Due to the absence of amino group at position 6, I base pairs with C residues instead of U residues, thus change the sequence of the cDNA synthesized by RT. Thus, one can detect unexpected A → G changes in the cDNA compared to the genomic sequence [95]. To confirm the data, it is also possible to compare the cDNA sequence obtained by direct sequencing of an unmodified transcript produced *by in vitro* transcription with the cDNA sequence obtained with the authentic RNA. Another recently developed method that is suitable for transcriptome-wide analysis of inosine is based on specific cyanoethylation of inosine [70].

5. Conclusion

Despite extensive development of specific chemical reagents during several decades (starting in the 1970s), the spectrum of available chemicals capable of selective reaction with modified ribonucleotides remains rather limited. This is related to a limited difference of chemical reactivity between modified nucleotides and their unmodified counterparts. Furthermore, the low stability of RNA molecules during required reaction conditions presents a stringent limitation. Due to these limitations, development of specific reaction schemes was possible only for a small subset of the known modified residues in RNA. Future development in the field should fill these gaps and propose new chemical reagents for extensive analysis of RNA modification by RT and by MS techniques.

Abbreviations

APM:	(*N*-acryloylamino)phenyl-4-mercuric chloride
APB:	(*N*-acryloylamino)phenyl -3-boronic acid
RT:	Reverse transcription
TLC:	Thin layer chromatography
CMCT:	1-cyclohexyl-3-(2-(4-morpholinyl)ethyl)*carbodiimide* tosylate
DMS:	Dimethylsulfate
DEPC:	Diethyl pyrocarbonate
DMAS-Cl:	N,N-(dimethylamino)dimethylchlorosilane
NHS:	N-hydroxysuccinimide.

References

[1] H. Grosjean and R. Benne, Eds., *Modification and Editing of RNA*, ASM Press, Washington, DC, USA, 1998.

[2] H. Grosjean, Ed., *Fine-Tuning of RNA Functions by Modification and Editing*, Springer, Berlin, Germany, 2005.

[3] H. Grosjean, Ed., *DNA and RNA Modification Enzymes: Structure, Mechanism, Function and Evolution*, Landes Bioscience, 2009.

[4] S. Douthwaite and F. Kirpekar, "Identifying modifications in RNA by MALDI mass spectrometry," *Methods in Enzymology*, vol. 425, pp. 1–20, 2007.

[5] T. S. Rozhdestvensky, P. F. Crain, and J. Brosius, "Isolation and posttranscriptional modification analysis of native BC1 RNA from mouse brain," *RNA Biology*, vol. 4, no. 1, pp. 11–15, 2007.

[6] H. Oberacher, "Frontiers of mass spectrometry in nucleic acids analysis," *European Journal of Mass Spectrometry*, vol. 16, no. 3, pp. 351–365, 2010.

[7] S. Kellner, J. Burhenne, and M. Helm, "Detection of RNA modifications," *RNA Biology*, vol. 7, no. 2, pp. 237–247, 2010.

[8] Y. Motorin, S. Muller, I. Behm-Ansmant, and C. Branlant, "Identification of modified residues in RNAs by reverse transcription-based methods," *Methods in Enzymology*, vol. 425, pp. 21–53, 2007.

[9] M. Silberklang, A. M. Gillum, and U. L. RajBhandary, "Use of in vitro 32P labeling in the sequence analysis of nonradioactive tRNAs," *Methods in Enzymology*, vol. 59, pp. 58–109, 1979.

[10] H. Grosjean, G. Keith, and L. Droogmans, "Detection and quantification of modified nucleotides in RNA using thin-layer chromatography," *Methods in Molecular Biology*, vol. 265, pp. 357–391, 2004.

[11] Z. Meng and P. A. Limbach, "Mass spectrometry of RNA: linking the genome to the proteome," *Briefings in Functional Genomics and Proteomics*, vol. 5, no. 1, pp. 87–95, 2006.

[12] Q. Dai, R. Fong, M. Saikia et al., "Identification of recognition residues for ligation-based detection and quantitation of pseudouridine and N-methyladenosine," *Nucleic Acids Research*, vol. 35, no. 18, pp. 6322–6329, 2007.

[13] A. M. Maxam and W. Gilbert, "A new method for sequencing DNA," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 74, no. 2, pp. 560–564, 1977.

[14] D. A. Peattie, "Direct chemical method for sequencing RNA," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 76, no. 4, pp. 1760–1764, 1979.

[15] D. A. Peattie and W. Gilbert, "Chemical probes for higher-order structure in RNA," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 77, no. 8, pp. 4679–4682, 1980.

[16] W. Wintermeyer and H. G. Zachau, "A specific chemical chain scission of tRNA at 7-methylguanosine," *FEBS Letters*, vol. 11, no. 3, pp. 160–164, 1970.

[17] R. Thiebe and H. G. Zachau, "A specific modification next to the anticodon of phenylalanine transfer ribonucleic acid," *European Journal of Biochemistry*, vol. 5, no. 4, pp. 546–555, 1968.

[18] B. Singer, "All oxygens in nucleic acids react with carcinogenic ethylating agents," *Nature*, vol. 264, no. 5584, pp. 333–339, 1976.

[19] S. A. Mortimer, J. S. Johnson, and K. M. Weeks, "Quantitative analysis of RNA solvent accessibility by N-silylation of guanosine," *Biochemistry*, vol. 48, no. 10, pp. 2109–2114, 2009.

[20] W. J. Krzyzosiak, J. Biernat, J. Ciesiołka, K. Gulewicz, and M. Wiewiórowski, "The reaction of adenine and cytosine residues in tRNA with chloroacetaldehyde," *Nucleic Acids Research*, vol. 9, no. 12, pp. 2841–2851, 1981.

[21] J. Biernat, J. Ciesiolka, and P. Gornicki, "New observations concerning the chloroacetaldehyde reaction with some tRNA constituents. Stable intermediates, kinetics and selectivity of the reaction," *Nucleic Acids Research*, vol. 5, no. 3, pp. 789–804, 1978.

[22] P. Roques, F. Thome, and M. Olomucki, "Reaction of nucleic acid bases with α-acetylenic esters. Part IV. Preparation of an alkylating derivative of tRNA(Phe) by conformation-specific chemical modification," *Biochimica et Biophysica Acta*, vol. 951, no. 1, pp. 71–77, 1988.

[23] K. Nakaya, O. Takenaka, H. Horinishi, and K. Shibata, "Reactions of glyoxal with nucleic acids, nucleotides and their component bases," *Biochimica et Biophysica Acta*, vol. 161, no. 1, pp. 23–31, 1968.

[24] H. F. Noller and J. B. Chaires, "Functional modification of 16S ribosomal RNA by kethoxal," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 69, no. 11, pp. 3113–3118, 1972.

[25] T. Pieler, A. Schreiber, and V. A. Erdmann, "Comparative structural analysis of eubacterial 5S rRNA by oxidation of adenines in the N-1 position," *Nucleic Acids Research*, vol. 12, no. 7, pp. 3115–3126, 1984.

[26] A. Okamoto, "Chemical approach toward efficient DNA methylation analysis," *Organic and Biomolecular Chemistry*, vol. 7, no. 1, pp. 21–26, 2009.

[27] I. Shcherbakova, S. Mitra, R. H. Beer, and M. Brenowitz, "Fast Fenton footprinting: a laboratory-based method for the time-resolved analysis of DNA, RNA and proteins," *Nucleic Acids Research*, vol. 34, no. 6, article no. e48, 2006.

[28] G. J. Murakawa, C. H. B. Chen, M. D. Kuwabara, D. P. Nierlich, and D. S. Sigman, "Scission of RNA by the chemical nuclease of 1,10-phenanthroline-copper ion: preference for single-stranded loops," *Nucleic Acids Research*, vol. 17, no. 13, pp. 5361–5375, 1989.

[29] Y. H. Wang, S. R. Sczekan, and E. C. Theil, "Structure of the 5' untranslated regulatory region of ferritin mRNA studied in solution," *Nucleic Acids Research*, vol. 18, no. 15, pp. 4463–4468, 1990.

[30] M. Götte, R. Marquet, C. Isel et al., "Probing the higher order structure of RNA with peroxonitrous acid," *FEBS Letters*, vol. 390, no. 2, pp. 226–228, 1996.

[31] Y. Kuchino, F. Mori, and H. Kasai, "Misreading of DNA templates containing 8-hydroxydeoxyguanosine at the modified base and at adjacent residues," *Nature*, vol. 326, no. 6117, pp. 77–79, 1987.

[32] D. Rhodes, "Accessible and inaccessible bases in yeast phenylalanine transfer RNA as studied by chemical modification," *Journal of Molecular Biology*, vol. 94, no. 3, pp. 449–460, 1975.

[33] K. Negishi, F. Harada, S. Nishimura, and H. Hayatsu, "A rapid cytosine-specific modification of E. coli tRNA (Leu) by semicarbazide-bisulfite, a probe for polynucleotide conformations," *Nucleic Acids Research*, vol. 4, no. 7, pp. 2283–2292, 1977.

[34] M. Münzel, L. Lercher, M. Müller, and T. Carell, "Chemical discrimination between dC and 5MedC via their hydroxylamine adducts," *Nucleic Acids Research*, vol. 38, no. 21, article e192, 2010.

[35] S. J. Clark, J. Harrison, C. L. Paul, and M. Frommer, "High sensitivity mapping of methylated cytosines," *Nucleic Acids Research*, vol. 22, no. 15, pp. 2990–2997, 1994.

[36] W. Gu, R. L. Hurto, A. K. Hopper, E. J. Grayhack, and E. M. Phizicky, "Depletion of *Saccharomyces cerevisiae* $tRNA^{His}$ guanylyltransferase Thg1p leads to uncharged $tRNA^{His}$ with additional m^5C," *Molecular and Cellular Biology*, vol. 25, no. 18, pp. 8191–8201, 2005.

[37] M. Schaefer, T. Pollex, K. Hanna, and F. Lyko, "RNA cytosine methylation analysis by bisulfite sequencing," *Nucleic Acids Research*, vol. 37, no. 2, article no. e12, 2009.

[38] I. L. Batey and D. M. Brown, "Chemical studies of the tertiary structure of transfer RNA. Iodination of exposed cytosine residues with I," *Biochimica et Biophysica Acta*, vol. 474, no. 3, pp. 378–385, 1977.

[39] J. L. Duh, R. R. Punzalan, and R. C. Pless, "Reactivity of pyrimidine nucleosides under the conditions of the pyrimidine sequencing reactions of Maxam and Gilbert," *Analytical Biochemistry*, vol. 168, no. 2, pp. 231–238, 1988.

[40] H. Hayatsu, "Reaction of cytidine with semicarbazide in the presence of bisulfite. A rapid modification specific for single-stranded polynucleotide," *Biochemistry*, vol. 15, no. 12, pp. 2677–2682, 1976.

[41] K. Miura, T. Iwano, S. Tsuda, T. Ueda, F. Harada, and N. Kato, "Chemical modification of cytosine residues of tRNAVa1 with hydrogen sulfide (nucleosides and nucleotides. XL)," *Chemical and Pharmaceutical Bulletin*, vol. 30, no. 11, pp. 4126–4133, 1982.

[42] N. Riehl, P. Carbon, B. Ehresmann, and J. P. Ebel, "Chemical conversion of cytidine residues into 4-thiouridines in yeast tRNAPhe. Determination of the modified cytidines," *Nucleic Acids Research*, vol. 12, no. 11, pp. 4445–4453, 1984.

[43] A. E. Johnson, H. J. Adkins, E. A. Matthews, and C. R. Cantor, "Distance moved by transfer RNA during translocation from the A site to the P site on the ribosome," *Journal of Molecular Biology*, vol. 156, no. 1, pp. 113–140, 1982.

[44] B. S. Watson, T. L. Hazlett, J. F. Eccleston, C. Davis, D. M. Jameson, and A. E. Johnson, "Macromolecular arrangement in the aminoacyl-tRNA·elongation factor Tu·GTP ternary complex. A fluorescence energy transfer study," *Biochemistry*, vol. 34, no. 24, pp. 7904–7912, 1995.

[45] P. Górnicki, M. Judek, A. Wolański, and W. J. Krzyzosiak, "Hypermodified nucleoside carboxyl group as a target site for specific tRNA modification," *European Journal of Biochemistry*, vol. 155, no. 2, pp. 371–375, 1986.

[46] J. Podkowinski, T. Dymarek-Babs, and P. Gornicki, "Modified tRNAs for probing tRNA binding sites on the ribosome," *Acta Biochimica Polonica*, vol. 36, no. 3-4, pp. 235–244, 1989.

[47] M. Schaefer, S. Hagemann, K. Hanna, and F. Lyko, "Azacytidine inhibits RNA methylation at DNMT2 target sites in human cancer cell lines," *Cancer Research*, vol. 69, no. 20, pp. 8127–8132, 2009.

[48] Y. Motorin, F. Lyko, and M. Helm, "5-methylcytosine in RNA: detection, enzymatic formation and biological functions," *Nucleic Acids Research*, vol. 38, no. 5, Article ID gkp1117, pp. 1415–1430, 2009.

[49] E. Fritzsche, H. Hayatsu, G. L. Igloi, S. Iida, and H. Kössel, "The use of permanganate as a sequencing reagent for identification of 5-methylcytosine residues in DNA," *Nucleic Acids Research*, vol. 15, no. 14, pp. 5517–5528, 1987.

[50] C. T. Bui, K. Rees, and R. G. H. Cotton, "Permanganate oxidation reactions of DNA: perspective in biological studies," *Nucleosides, Nucleotides and Nucleic Acids*, vol. 22, no. 9, pp. 1835–1855, 2003.

[51] K. Tanaka, K. Tainaka, T. Kamei, and A. Okamoto, "Direct labeling of 5-methylcytosine and its applications," *Journal of the American Chemical Society*, vol. 129, no. 17, pp. 5612–5620, 2007.

[52] K. Tanaka, K. Tainaka, and A. Okamoto, "Methylcytosine-selective fluorescence quenching by osmium complexation," *Bioorganic and Medicinal Chemistry*, vol. 15, no. 4, pp. 1615–1621, 2007.

[53] K. Tanaka, K. Tainaka, T. Umemoto, A. Nomura, and A. Okamoto, "An osmium-DNA interstrand complex: application to facile DNA methylation analysis," *Journal of the American Chemical Society*, vol. 129, no. 46, pp. 14511–14517, 2007.

[54] W. Wintermeyer and H. G. Zachau, "Tertiary structure interactions of 7 methylguanosine in yeast tRNA(Phe) as studied by borohydride reduction," *FEBS Letters*, vol. 58, no. 1, pp. 306–309, 1975.

[55] B. E. H. Maden, M. E. Corbett, P. A. Heeney, K. Pugh, and P. M. Ajuh, "Classical and novel approaches to the defection and localization of the numerous modified nucleotides in eukaryotic ribosomal RNA," *Biochimie*, vol. 77, no. 1-2, pp. 22–29, 1995.

[56] B. E. H. Maden, "Mapping 2'-O-methyl groups in ribosomal RNA," *Methods*, vol. 25, no. 3, pp. 374–382, 2001.

[57] Y. T. Yu, M. D. Shu, and J. A. Steitz, "A new method for detecting sites of 2'-O-methylation in RNA molecules," *RNA*, vol. 3, no. 3, pp. 324–331, 1997.

[58] B. Yu, Z. Yang, J. Li et al., "Methylation as a crucial step in plant microRNA biogenesis," *Science*, vol. 307, no. 5711, pp. 932–935, 2005.

[59] B. Yu, E. J. Chapman, Z. Yang, J. C. Carrington, and X. Chen, "Transgenically expressed viral RNA silencing suppressors interfere with microRNA methylation in Arabidopsis," *FEBS Letters*, vol. 580, no. 13, pp. 3117–3120, 2006.

[60] S. Alefelder, B. K. Patel, and F. Eckstein, "Incorporation of terminal phosphorothioates into oligonucleotides," *Nucleic Acids Research*, vol. 26, no. 21, pp. 4983–4988, 1998.

[61] E. J. Merino, K. A. Wilkinson, J. L. Coughlan, and K. M. Weeks, "RNA structure analysis at single nucleotide resolution by Selective 2'-Hydroxyl Acylation and Primer Extension (SHAPE)," *Journal of the American Chemical Society*, vol. 127, no. 12, pp. 4223–4231, 2005.

[62] K. A. Wilkinson, E. J. Merino, and K. M. Weeks, "Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution," *Nature Protocols*, vol. 1, no. 3, pp. 1610–1616, 2006.

[63] R. P. Singhal, R. K. Bajaj, and C. M. Buess, "Reversed-phase boronate chromatography for the separation of O-methylribose nucleosides and aminoacyl-tRNAs," *Analytical Biochemistry*, vol. 109, no. 1, pp. 1–11, 1980.

[64] G. L. Igloi and H. Kössel, "Affinity electrophoresis for monitoring terminal phosphorylation and the presence of queuosine in RNA. Application of polyacrylamide containing a covalently bound boronic acid," *Nucleic Acids Research*, vol. 13, no. 19, pp. 6881–6898, 1985.

[65] A. Durairaj and P. A. Limbach, "Matrix-assisted laser desorption/ionization mass spectrometry screening for pseudouridine in mixtures of small RNAs by chemical derivatization, RNase digestion and signature products," *Rapid Communications in Mass Spectrometry*, vol. 22, no. 23, pp. 3727–3734, 2008.

[66] A. Durairaj and P. A. Limbach, "Mass spectrometry of the fifth nucleoside: a review of the identification of pseudouridine in nucleic acids," *Analytica Chimica Acta*, vol. 623, no. 2, pp. 117–125, 2008.

[67] A. Durairaj and P. A. Limbach, "Improving CMC-derivatization of pseudouridine in RNA for mass spectrometric detection," *Analytica Chimica Acta*, vol. 612, no. 2, pp. 173–181, 2008.

[68] J. Mengel-Jørgensen and F. Kirpekar, "Detection of pseudouridine and other modifications in tRNA by cyanoethylation and MALDI mass spectrometry," *Nucleic Acids Research*, vol. 30, no. 23, article e135, 2002.

[69] G. Emmerechts, P. Herdewijn, and J. Rozenski, "Pseudouridine detection improvement by derivatization with methyl vinyl sulfone and capillary HPLC-mass spectrometry," *Journal of Chromatography B*, vol. 825, no. 2, pp. 233–238, 2005.

[70] M. Sakurai, T. Yano, H. Kawabata, H. Ueda, and T. Suzuki, "Inosine cyanoethylation identifies A-to-I RNA editing sites in the human transcriptome," *Nature Chemical Biology*, vol. 6, no. 10, pp. 733–740, 2010.

[71] T. Mizutani, "Chemical modification of pseudouridylic acid in valine transfer ribonucleic acid I from Torulopsis utilis with a radioactive carbodiimide," *Journal of Biochemistry*, vol. 71, no. 4, pp. 589–595, 1972.

[72] T. Mizutani, "Chemical modification of pseudouridylic acid in isoeleucine transfer ribonucleic acid from Torulopsis utilis with a radioactive carbodiimide," *Journal of Biochemistry*, vol. 71, no. 3, pp. 567–569, 1972.

[73] N. W. Y. Ho and P. T. Gilham, "Reaction of pseudouridine and inosine with N-cyclohexyl-N'-β-(4-methylmorpholinium) ethylcarbodiimide," *Biochemistry*, vol. 10, no. 20, pp. 3651–3657, 1971.

[74] K. G. Patteson, L. P. Rodicio, and P. A. Limbach, "Identification of the mass-silent post-transcriptionally modified nucleoside pseudouridine in RNA by matrix-assisted laser desorption/ionization mass spectrometry," *Nucleic acids research*, vol. 29, no. 10, p. E49, 2001.

[75] A. Bakin and J. Ofengand, "Four newly located pseudouridylate residues in Escherichia coli 23S ribosomal RNA are all at the peptidyltransferase center: analysis by the application of a new sequencing technique," *Biochemistry*, vol. 32, no. 37, pp. 9754–9762, 1993.

[76] A. Bakin and J. Ofengand, "Mapping of pseudouridine residues in RNA to nucleotide resolution," in *Protein Synthesis: Methods and Protocols*, R. Martin, Ed., vol. 77, pp. 297–309, Springer, Berlin, Germany, 1998.

[77] S. Massenet, Y. Motorin, D. L. J. Lafontaine, E. C. Hurt, H. Grosjean, and C. Branlant, "Pseudouridine mapping in the Saccharomyces cerevisiae spliceosomal U small nuclear RNAs (snRNAs) reveals that pseudouridine synthase Pus1p exhibits a dual substrate specificity for U2 snRNA and tRNA," *Molecular and Cellular Biology*, vol. 19, no. 3, pp. 2142–2154, 1999.

[78] M. Caron and H. Dugas, "Specific spin labeling of transfer ribonucleic acid molecules," *Nucleic Acids Research*, vol. 3, no. 1, pp. 19–34, 1976.

[79] G. L. Igloi, "Interaction of tRNAs and of phosphorothioate-substituted nucleic acids with an organomercurial. Probing the chemical environment of thiolated residues by affinity electrophoresis," *Biochemistry*, vol. 27, no. 10, pp. 3842–3849, 1988.

[80] S. Friedman, H. J. Li, K. Nakanishi, and G. Van Lear, "3-(3-Amino-3-carboxy-n-propyl)uridine. the structure of the nucleoside in Escherichia coli transfer ribonucleic acid that reacts with phenoxyacetoxysuccinimide," *Biochemistry*, vol. 13, no. 14, pp. 2932–2937, 1974.

[81] J. A. Plumbridge, H. G. Baumert, M. Ehrenberg, and R. Rigler, "Characterisation of a new, fully active fluorescent derivative of E. coli tRNA Phe," *Nucleic Acids Research*, vol. 8, no. 4, pp. 827–843, 1980.

[82] A. Pingoud, R. Kownatzki, and G. Maass, "Fluoresceinylthiocarbamyl tRNA(Tyr): a useful derivative of tRNA(Tyr) (E. coli) for physicochemical studies," *Nucleic Acids Research*, vol. 4, no. 2, pp. 327–338, 1977.

[83] H. Kasai, N. Shindo-Okada, S. Noguchi, and S. Nishimura, "Specific fluorescent labelling of 7-(aminomethyl)-7-deazaguanosine located in the anticodon of tRNA(Tyr) isolated from E. coli mutant," *Nucleic Acids Research*, vol. 7, no. 1, pp. 231–238, 1979.

[84] P. Cerutti and N. Miller, "Selective reduction of yeast transfer ribonucleic acid with sodium borohydride," *Journal of Molecular Biology*, vol. 26, no. 1, pp. 55–66, 1967.

[85] T. Igo-Kemenes and H. G. Zachau, "On the specificity of the reduction of transfer ribonucleic acids with sodium borohydride," *European Journal of Biochemistry*, vol. 10, no. 3, pp. 549–556, 1969.

[86] W. Wintermeyer, H. G. Schleich, and H. G. Zachau, "Incorporation of amines or hydrazines into tRNA replacing wybutine or dihydrouracil," *Methods in Enzymology*, vol. 59, pp. 110–121, 1979.

[87] F. Xing, S. L. Hiley, T. R. Hughes, and E. M. Phizicky, "The specificities of four yeast dihydrouridine synthases for cytoplasmic tRNAs," *Journal of Biological Chemistry*, vol. 279, no. 17, pp. 17850–17860, 2004.

[88] C. H. House and S. L. Miller, "Hydrolysis of dihydrouridine and related compounds," *Biochemistry*, vol. 35, no. 1, pp. 315–320, 1996.

[89] B. L. Bass, "An I for editing," *Current Biology*, vol. 5, no. 6, pp. 598–600, 1995.

[90] S. Maas, A. Rich, and K. Nishikura, "A-to-I RNA editing: recent news and residual mysteries," *Journal of Biological Chemistry*, vol. 278, no. 3, pp. 1391–1394, 2003.

[91] D. P. Morse and B. L. Bass, "Detection of inosine in messenger RNA by inosine-specific cleavage," *Biochemistry*, vol. 36, no. 28, pp. 8429–8434, 1997.

[92] G. Vogeli, T. S. Stewart, T. McCutchan, and D. Soll, "Isolation of Escherichia coli precursor tRNAs containing modified nucleoside Q," *Journal of Biological Chemistry*, vol. 252, no. 7, pp. 2311–2318, 1977.

[93] J. E. Ladner and M. P. Schweizer, "Effects of dilute HCl on yeast tRNAPhe and E. coli tRNA1fMet," *Nucleic Acids Research*, vol. 1, no. 2, pp. 183–192, 1974.

[94] H. G. Schleich, W. Wintermeyer, and H. G. Zachau, "Replacement of wybutine by hydrazines and its effect on the active conformation of yeast tRNA(Phe)," *Nucleic Acids Research*, vol. 5, no. 5, pp. 1701–1713, 1978.

[95] L. Saccomanno and B. L. Bass, "A minor fraction of basic fibroblast growth factor mRNA is deaminated in Xenopus stage VI and matured oocytes," *RNA*, vol. 5, no. 1, pp. 39–48, 1999.

SAGE-Hindawi Access to Research
Journal of Nucleic Acids
Volume 2011, Article ID 586935, 7 pages
doi:10.4061/2011/586935

Research Article

Branched RNA: A New Architecture for RNA Interference

Anna Aviñó,[1] Sandra M. Ocampo,[1] José Carlos Perales,[2] and Ramon Eritja[1]

[1] *Institute for Research in Biomedicine (IRB Barcelona), Institute of Advanced Chemistry of Catalonia (IQAC-CSIC), Networking Centre on Bioengineering, Biomaterials, and Nanomedicine (CIBER-BBN), Baldiri Reixac 10, 08028 Barcelona, Spain*
[2] *Department of Physiological Sciences, School of Medicine, University of Barcelona (Campus Bellvitge), Feixa Llarga, L'Hospitalet de Llobregat, 08907 Barcelona, Spain*

Correspondence should be addressed to Ramon Eritja, recgma@cid.csic.es

Received 16 October 2010; Accepted 14 January 2011

Academic Editor: Arthur van Aerschot

Branched RNAs with two and four strands were synthesized. These structures were used to obtain branched siRNA. The branched siRNA duplexes had similar inhibitory capacity as those of unmodified siRNA duplexes, as deduced from gene silencing experiments of the TNF-α protein. Branched RNAs are considered novel structures for siRNA technology, and they provide an innovative tool for specific gene inhibition. As the method described here is compatible with most RNA modifications described to date, these compounds may be further functionalized to obtain more potent siRNA derivatives and can be attached to suitable delivery systems.

1. Introduction

In recent years, siRNAs have generated tremendous interest in therapeutics [1]. Nevertheless, the transition of siRNAs from the laboratory to the clinical practice has encountered several obstacles. Briefly, siRNA duplexes are rapidly degraded in serum by exonucleases and endonucleases [2]. The polyanionic phosphodiester backbone of siRNA suffers from difficult cell uptake [3], and oligonucleotides may have off-target effects, either by stimulating the immune system [4] or by entering other endogenous gene regulation pathways [5]. Several chemical modifications have been proposed in the literature to address these drawbacks [2–4]. Most of these modifications are based on modified nucleosides and changes on backbone linkages [6, 7]. Thus, changes in sugar moiety influences sugar conformation, and, therefore, overall siRNA structure. Modifications of the 2′-OH by F or OMe as well as LNA [8, 9] are well tolerated and improve binding affinity and nuclease resistance. Base modifications that stabilize base pairs (5-bromouracil, 5-methylcytosine, 5-propynyluracil, and others) have also been proposed [7, 10]. Terminal conjugates, especially at the termini of the sense strand, have been modified with a large number of lipids to achieve improved cellular uptake [11].

In addition to these modifications, siRNA architecture is also crucial in the design of effective and specific siRNA. The architecture itself can be altered by chemical synthesis. In addition to the canonical siRNA architecture of 21-nt antiparallel, double-strand RNA with 2-nt 3′-overhangs [12], several forms of siRNA have been described. Blunt-ended siRNA [13], 25/27 mer Dicer-substrate or asymmetric siRNA [14] are among the siRNA structures formed by two strands. Moreover, functional siRNA can also be formed by one single RNA strand. This is the case in small hairpin RNA (shRNA), where the two strands are linked by a single loop [15], or RNA dumbbells [16], made by closing the open end of the hairpin. This last structure retains RNAi activity while providing complete protection from nucleases [16]. Finally, siRNA can also comprise three strands, namely, two 9–13 nt sense strands and the intact antisense strand. This structure is known as small internally segmented interfering RNA [17] (sisiRNA). Some of these modifications have reduced off-target effects and increased potency (Figure 1(a)). Another architecture not yet explored in siRNA is the branched RNA structure obtained from a central building unit and several branching points that enable the strand growth.

Several strategies can be used to prepare branched RNA structures. Although the synthesis of these compounds is

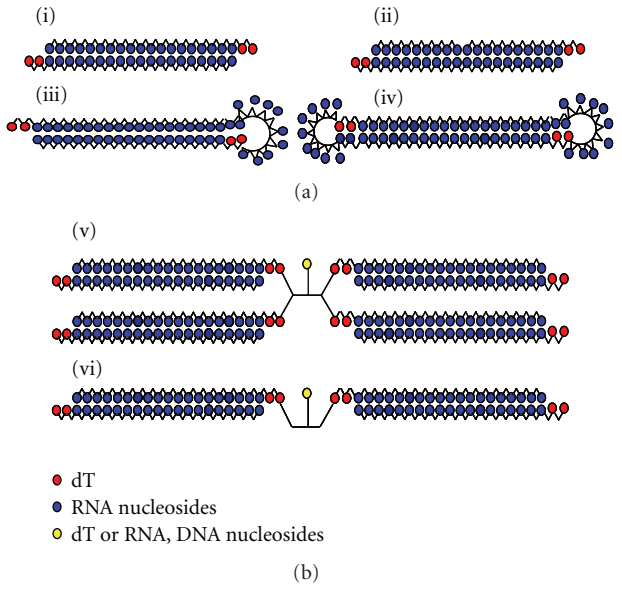


Figure 1: Duplex RNA architectures for RNA interference. (a) Described previously in the introduction: (i) canonical siRNA antiparallel duplex; (ii) small internally segmented interfering RNA (sisiRNA); (iii) small hairpin RNA (shRNA); (iv) dumbbell siRNA. (b) Branched siRNA described in this study: (v) four-stranded RNA; (vi) two-stranded RNA.

complex and tedious, commercially available synthons have improved the complexity and yields of these structures.

The assembly of branched nucleic acids on a solid support can be achieved by convergent or divergent strategies. In the former, synthesis of branched oligonucleotides containing two or more identical strands can be achieved by branching derivatives 1,3-diaminopropanol, pentaerythritol, the commercially available symmetric doubler [18–20], or by a ribonucleoside bisphosphoramidite [21] as synthons. In contrast, in the divergent approach two or more distinct strands are prepared with synthons with orthogonal protecting groups or from commercial sources [20, 22]. In addition, 2′-O-silylribonucleosides have been used to synthesize asymmetric double oligonucleotide strands [23].

Synthetic branched oligonucleotides have been applied for several purposes. Initially, most of the interest in this area was focused on the study of branched oligoribonucleotides as splicing intermediates of eukaryotic mRNAs [24–26]. Moreover, branched oligonucleotides show high affinity for single-strand oligonucleotides to form alternated strand triplexes [27–29]. Recently, branched oligonucleotides have been used as building blocks in the synthesis of new nanostructures [30–34]. Multilabelled oligonucleotides containing branching points have been described to increase the sensitivity of hybridization experiments [35].

Here, we synthesized branched RNA structures (Figure 1(b)) and evaluated their capacity to inhibit the tumour necrosis factor (TNF-α) protein, which is involved in the apoptosis, inflammation, and immunity processes [36]. We reasoned that branched siRNA could provide a new RNA architecture for RNA interference activity. Given that the use of symmetric branching units is compatible with most of the modifications described to enhance the inhibitory capacity of siRNA, the molecules described here may provide a starting point for further modifications.

2. Experimental Section

2.1. Oligonucleotides. The following RNA sequences were obtained from commercial sources (*Sigma-Proligo*, *Dharmacon*): sense or passenger scrambled 5′-CAGUCGCGUUUGCGACUGG-dT-dT-3′, antisense or guide scrambled 5′-CCAGUCGCAAACGCGACUG-dT-dT-3′, antisense or guide anti-TNF-α: 5′-GAGGCUGAGACAUAGGCAC-dT-dT-3′, and sense or passenger anti-TNF-α: 5′-GUGCCUAUGUCUCAGCCUC-dT-dT-3′. RNA monomers in capital letters, dT represents thymidine. The anti-TNFα siRNA was previously described to efficiently downregulate murine TNFα mRNA [36].

Figure 2 refers to the branched RNA structures synthesized in this study. DB stands for the symmetric doubler phosphoramidite obtained from commercial sources (*Glen Research*). Guanosine was protected with the dimethylaminomethylidene group, cytidine with the acetyl group, and adenosine with the benzoyl group. *t*-Butyldimethylsilyl (TBDMS) group was used for the protection of the 2′-OH function of the RNA monomers. The phosphoramidites were dissolved in dry acetonitrile (0.1 M), and a modified cycle was used with increased coupling time to 10 min. Oligoribonucleotide **1** was synthesized on a CPG solid support with a symmetric branching unit of two arms containing two DMT-protected hydroxyl groups, as described in [31]. Oligoribonucleotides **2** and **3** were synthesized using standard low-volume polystyrene thymidine columns. After the solid-phase synthesis, the supports were treated with concentrated aqueous ammonia-ethanol (3 : 1) for 1 h at 55°C. After filtration of the supports, the solutions were evaporated to dryness. The residue was dissolved in 85 μL of 1 M tetrabutylammonium fluoride (TBAF) in tetrahydrofuran (THF) for 12 h. Then, 85 μL of 1 M of triethylammonium acetate was added and the oligoribonucleotides were desalted on a NAP-10 column using water as eluent. The compounds were purified by HPLC under the following conditions. Column: Nucleosil 120–10 C_{18} (250 × 4 mm); 20-min linear gradient from 15% to 100% B (DMT ON conditions); flow rate 3 mL/min; solution A was 5% acetonitrile in 0.1 M aqueous triethylammonium acetate (TEAA) buffer and B 70% acetonitrile in 0.1 M aqueous TEAA. The purified products were analyzed by MALDI-TOF mass spectrometry. Yields (0.2 μmol scale synthesis) were between 5–10 OD units at 260 nm.

2.2. Thermal Denaturation Studies. The thermal melting curves for duplexes of the oligoribonucleotides **1–3** and their unmodified RNA complementary strands (guide strand) were performed following the absorption change at 260 nm. Samples were heated from 20°C to 80°C, with a linear temperature ramp of 0.5°/min in a JASCO V-650 spectrophotometer equipped with a Peltier temperature control. Sample concentration of the samples was around 2 μM. All the

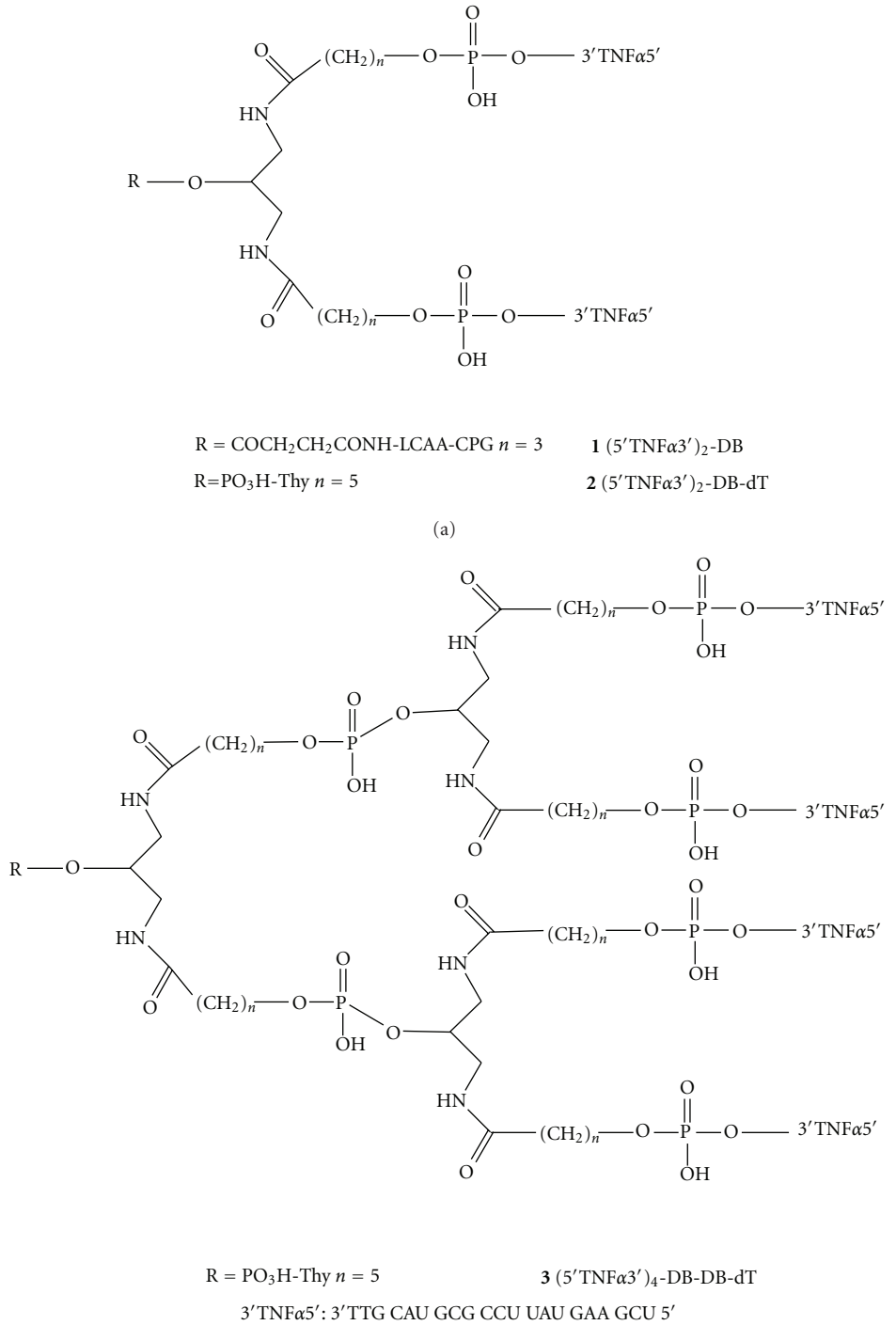


Figure 2: Schematic representation of the chemical structure of the branching units of the oligonucleotides described in this study.

measurements were repeated three times and conducted in 15 mM HEPES 1 mM $Mg(OAc)_2$ and 50 mM KOAc pH 7.4.

2.3. Cell Culture, Transfection, and Cellular Assays. HeLa cells were cultured under standard conditions (37°C, 5% CO_2, Dulbecco's Modified Eagle Medium, 10% fetal bovine serum, 2 mM L-glutamine, supplemented with penicillin (100 U/mL) and streptomycin (100 mg/mL)). All *in vitro* experiments were performed at 40–60% confluence. HeLa cells were transfected with 250 ng of a plasmid expressing murine TNF-α using lipofectin (*Invitrogen*), following the manufacturer's instructions. One hour after transfection cells were transfected with 100 nM double strand concentration of siRNA against TNF-α, using oligofectamine (*Invitrogen*).

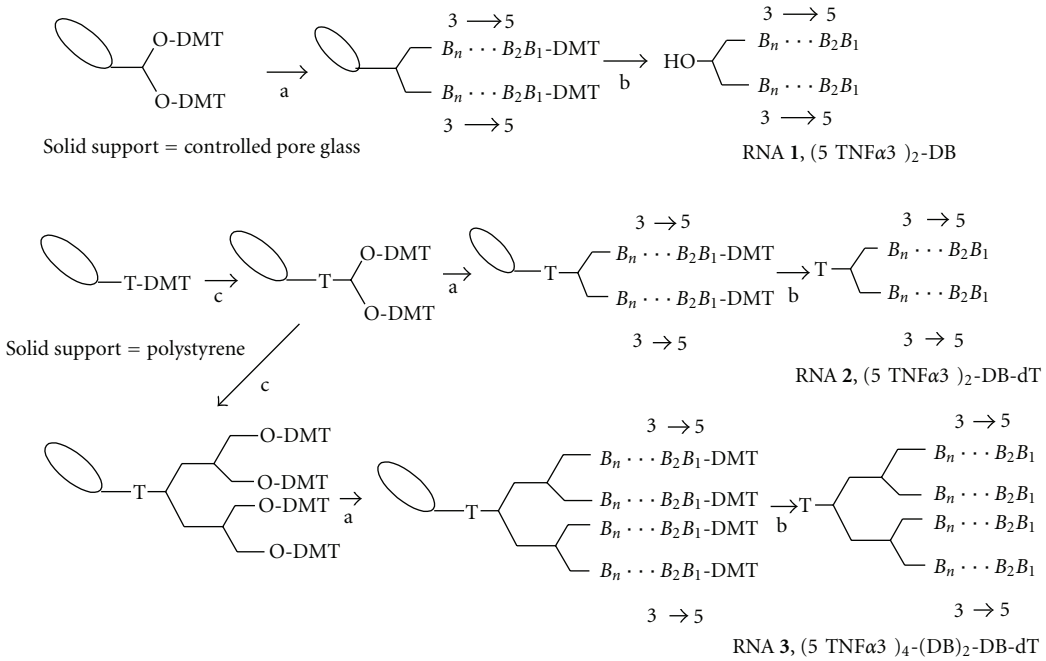


Figure 3: Outline of the synthesis of branched RNA oligonucleotides **1**, **2**, and **3** with two and four arms with identical sequence. (a) Assembly of the RNA sequence by standard solid-phase methods; (b) removal of protecting groups and release of the RNA molecule from solid support; (c) addition of the symmetrical branching phosphoramidite.

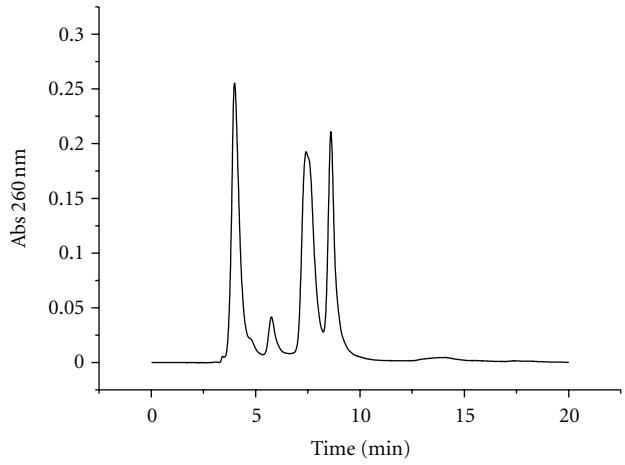


Figure 4: HPLC profile of DMT-containing oligonucleotide **2** with two arms. Truncated sequences without DMT groups had a retention of less than 5 min. Fraction eluting between 7-8 min contained oligonucleotides with a single DMT group. The last fraction contained the desired sequence with two DMT groups.

Previously, siRNA duplex annealing was performed by mixing modified (**1**, **2**, **3**) and unmodified passenger strands (**unm**) with the appropriate amount of the corresponding unmodified guide strand.

TNF-α concentration was determined from cell culture supernatant by enzyme-linked immunosorbent assay kit (*Bender MedSystems*) following the manufacturer's instructions. The inhibitory capacity of the siRNA duplexes is expressed as double strand concentration for comparative purposes. A 100 mM double strand concentration is equivalent to a 50 nM concentration of two-branched siRNA (**1** or **2**) and to 25 nM of four-branched siRNA (**3**).

3. Results and Discussion

3.1. Oligonucleotide Synthesis. In order to prepare branched RNA for RNA interference, the potential steric hindrance of the branching unit with RISC must be considered. As the passenger strand is removed from the siRNA duplex upon binding to RISC, we introduced the branching modification at the protruding 3′-end of the sense strand. This position has been demonstrated to allow the introduction of a large number of modifications without affecting the inhibitory capacity of siRNA [6, 7]. We thus designed branched oligonucleotide sequences **1–3** of the passenger strand of a siRNA directed against TNF-α (Figure 2). Sequence **1** was synthesized using a controlled pore glass (CPG) solid support containing a symmetric doubler [31], as shown in Figure 3. Sequences **2** and **3** with two or four strands, respectively, were synthesized on a low-volume polystyrene support (LV200) functionalized with dimethoxytrityl- (DMT-) thymidine. The commercially available symmetric doubler phosphoramidite was used to introduce two and four branches on the 3′-position of the starting thymidine (Figure 3). Sequences were assembled using standard protocols for RNA synthesis. The 2′-OH function of ribonucleosides was protected with the *t*-butyldimethylsilyl (TBDMS) group. Coupling yields, determined by the absorbance of the DMT

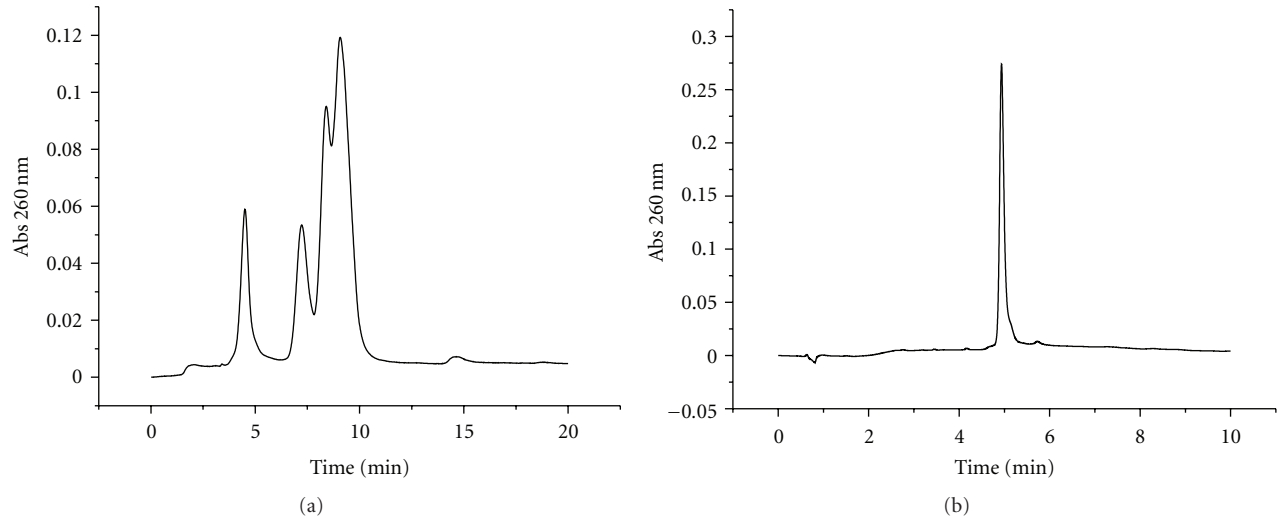


Figure 5: HPLC purification of oligonucleotide **3** with four arms. (a) HPLC profile of DMT-containing oligonucleotide **3**. The last fraction contained the desired sequence with four DMT groups; (b) analytical HPLC of purified oligonucleotide **3** after removal of DMT groups.

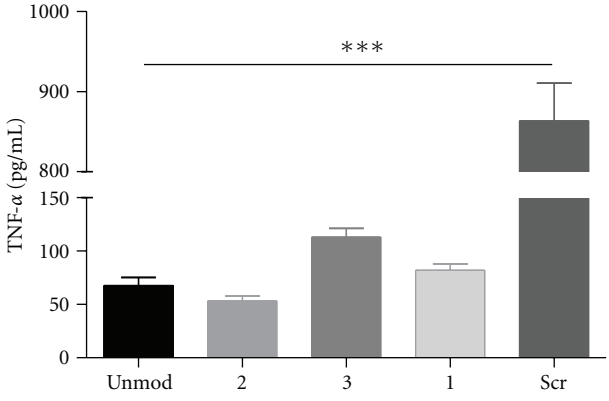


Figure 6: Inhibitory capacity of branched anti-TNF-α siRNAs. The inhibitory capacity of the siRNA duplexes are expressed as double strand concentration for comparative purposes. A 100-nM double strand concentration is equivalent to a concentration of 100 nM of unmodified siRNA duplex, 50 nM of siRNA **1**, and **2** and 25 nM of siRNA **3**. Transfection of siRNAs was carried out using oligofectamine. Values are represented as the average ±ES, $n = 3$ and are compared to a scrambled sequence. $^{***}P < .001$, ANOVA Test, Bonferroni post-test.

cation released in each synthesis step, were more efficient (98%) on low-volume (LV200) polystyrene supports than on CPG support (95%). After assembly of the sequences, the DMT-containing oligonucleotides were released from the supports with ammonia, and the resulting compounds were treated with fluoride to remove the TBDMS groups. HPLC analysis of the resulting products is shown in Figures 4 and 5. Several peaks were observed for the synthesis of the two-branch RNA sequences (**1** and **2**; Figure 4). Truncated sequences without DMT groups eluted between 3–5 min. A fraction containing oligonucleotides with a single DMT group was eluted next, and the last fraction contained the desired sequence with two DMT groups. Mass spectrometry analysis (Table 1) and electrophoresis analysis confirmed the mass and size of the desired branched oligoribonucleotides.

Figure 5 shows the HPLC profile of the mixture obtained in the synthesis of the four-branch RNA sequence (**3**). In this case, three peaks in the DMT-containing area were observed. Although resolution of these peaks was not as good as in the previous case, the last eluting peak corresponded to the desired tetra-DMT compound (**3**). The purified compound had the expected molecular weight, was homogeneous by analytical HPLC (Figure 5), and showed the correct migration in polyacrylamide gel electrophoresis (PAGE).

3.2. Thermal Denaturation Studies. The melting temperatures of the branched siRNA duplexes formed by annealing of equimolar amounts of sequences **1–3** with unmodified passenger strand are shown in Table 1. Duplex **1** had the lowest melting temperature, which was 3.5°C lower than the unmodified duplex (Table 1). Duplex **2** melted 1.5°C lower than the unmodified duplex. In contrast, duplex **3** had similar melting temperatures as the unmodified duplex. The small decrease in melting temperatures of the two-branched siRNA structures is possibly due to a steric effect in the branching point that holds the two duplex strands in close proximity. The four-stranded architecture had a larger separation between strands as a result of the introduction of 3 branching units, thus the resulting duplexes showed greater similarity to the unmodified duplex. Thus we believe that the small destabilizing effect observed in the two-branched RNA duplexes could be optimized in further experiments by adding a linker between the branching unit and the RNA strands, as described by Grimau et al. [31].

3.3. Cell Culture, Transfection, and Cellular Assays.. Tumor necrosis factor (TNF-α) was selected as a target for RNA interference studies. This protein is a major mediator of apoptosis as well as inflammation and immunity, and it has

Table 1: Mass spectrometry data on modified passenger strand and melting temperatures of siRNA duplexes formed by oligonucleotides **1**, **2**, and **3** and the linear unmodified control sequences. Buffer conditions: 15 mM HEPES, 1 mM magnesium acetate, 50 mM potassium acetate pH 7.4.

N°	oligonucleotides	MS (expected)	MS (found)	Tm (°C)	ΔT (°C)
1	$(5'\text{TNF}\alpha\ 3')_2$-DB	13508	13504	79.8	−3.5
2	$(5'\text{TNF}\alpha\ 3')_2$-DB-dT	13842	13847	82.5	−1.7
3	$(5'\text{TNF}\alpha\ 3')_4$-$(\text{DB})_2$-DB-dT	27798	27823	84.3	+0.1
Unmodified	Passenger TNFα	n.d.	n.d.	84.2	0

been implicated in the pathogenesis of a wide spectrum of human diseases. Consequently, inhibition of this protein is of particular relevance. Modified oligoribonucleotides (**1**–**3**) were annealed with equimolar amounts of the unmodified guide, and the resulting duplexes were used to inhibit the expression of TNF-α gene. HeLa cells were transfected first with the murine TNF-α plasmid using lipofectin, and 1 h later they were cotransfected with the siRNA duplex using oligofectamine. After 24 h, cellular TNF-α production was analyzed by enzyme-linked immunosorbent assay (ELISA). The inhibitory capacity of the siRNA duplexes is shown in Figure 6. To compare the efficiency of each siRNA to inhibit TNF-α, we normalized the data taking in account the number of strands of each siRNA. Thus, a 100 nM double strand concentration is equivalent to 100 nM of unmodified siRNA duplex, 50 nM of siRNA **1**, and **2**, and 25 nM of siRNA **3**. Figure 6 shows that the inhibitory capacity of the branched structures was maintained similar to that of the unmodified duplex. This result indicates that the branched siRNAs described here are compatible with RNA interference machinery, and thus the RISC complex binds to branched RNA structures in a similar way as to the linear RNA duplexes shown in Figure 1. Two-stranded RNA duplexes (**1** and **2**) were more efficient than four-stranded ones (**3**). In addition, siRNA **1** showed slightly greater efficiency at inhibiting TNF-α than siRNA duplex **2**. This small difference may be related to the lower melting temperature of the former (Table 1).

4. Conclusions

For several years, research has focused on chemical modifications and delivery technologies to improve the pharmacokinetic properties of siRNA. Many of the chemically modified siRNA with interesting inhibitory capacity contain one or multiple modifications in the sugar, nucleobases, and phosphate linkages or at the 3′- or 5′-ends. In addition to these modifications, duplex architecture of siRNA itself is also relevant, and several modifications have been reported to show satisfactory inhibitory capacity. Here we demonstrate that branched siRNA is compatible with RNAi and that, when transfected with cationic lipids, siRNA has similar inhibitory capacity than unmodified duplex siRNA. Although the potency of branched siRNA containing two or four strands was not increased, we consider it a suitable starting point for further development. Given that the method described here is compatible with most of the RNA modifications described to date, these compounds may be further functionalized to obtain more potent siRNA derivatives. In addition, they offer an internal mid position that could be suitable for attachment to delivery systems. In this regard, optimization of the branching approach for the synthesis of asymmetric branched siRNAs may lead to the development of siRNA for the combined inhibition of multiple targets. These asymmetric siRNA duplexes carrying two RNA sequences attached or bound to an appropriate delivery system will insure the 1 : 1 ratio of two RNA sequences for the combined inhibition of two genes that may improve the treatment of a particular disease.

Acknowledgments

This paper was supported by the Spanish Ministry of Education (Grants BFU2007-63287, CTQ2010-20541) and the *Generalitat de Catalunya* (2009/SGR/208). CIBER-BBN is an initiative funded by the VI National R&D&i Plan *2008–2011, Iniciativa Ingenio 2010, Consolider Program, and CIBER Actions* and financed by the *Instituto de Salud Carlos III* with assistance from the *European Regional Development Fund.*

References

[1] A. de Fougerolles, H. P. Vornlocher, J. Maraganore, and J. Lieberman, "Interfering with disease: a progress report on siRNA-based therapeutics," *Nature Reviews Drug Discovery*, vol. 6, no. 6, pp. 443–453, 2007.

[2] D. A. Braasch, S. Jensen, Y. Liu et al., "RNA interference in mammalian cells by chemically-modified RNA," *Biochemistry*, vol. 42, no. 26, pp. 7967–7975, 2003.

[3] K. Tiemann and J. J. Rossi, "RNAi-based therapeutics-current status, challenges and prospects," *EMBO Molecular Medicine*, vol. 1, no. 3, pp. 142–151, 2009.

[4] F. Eberle, K. Gießler, C. Deck et al., "Modifications in small interfering RNA that separate immunostimulation from RNA interference," *Journal of Immunology*, vol. 180, no. 5, pp. 3229–3237, 2008.

[5] A. L. Jackson, S. R. Bartz, J. Schelter et al., "Expression profiling reveals off-target gene regulation by RNAi," *Nature Biotechnology*, vol. 21, no. 6, pp. 635–637, 2003.

[6] J. K. Watts, G. F. Deleavey, and M. J. Damha, "Chemically modified siRNA: tools and applications," *Drug Discovery Today*, vol. 13, no. 19-20, pp. 842–855, 2008.

[7] YA. L. Chiu and T. M. Rana, "siRNA function in RNAi: a chemical modification analysis," *RNA*, vol. 9, no. 9, pp. 1034–1048, 2003.

[8] R. A. Blidner, R. P. Hammer, M. J. Lopez, S. O. Robinson, and W. T. Monroe, "Fully 2′-deoxy-2′-fluoro substituted nucleic acids induce RNA interference in mammalian cell culture,"

Chemical Biology and Drug Design, vol. 70, no. 2, pp. 113–122, 2007.

[9] J. Elmén, H. Thonberg, K. Ljungberg et al., "Locked nucleic acid (LNA) mediated improvements in siRNA stability and functionality," *Nucleic Acids Research*, vol. 33, no. 1, pp. 439–447, 2005.

[10] M. Terrazas and E. T. Kool, "RNA major groove modifications improve siRNA stability and biological activity," *Nucleic Acids Research*, vol. 37, no. 2, pp. 346–353, 2009.

[11] J. Soutschek, A. Akinc, B. Bramlage et al., "Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs," *Nature*, vol. 432, no. 7014, pp. 173–178, 2004.

[12] S. M. Elbashir, J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, and T. Tuschl, "Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells," *Nature*, vol. 411, no. 6836, pp. 494–498, 2001.

[13] F. Czauderna, M. Fechtner, S. Dames et al., "Structural variations and stabilising modifications of synthetic siRNAs in mammalian cells," *Nucleic Acids Research*, vol. 31, no. 11, pp. 2705–2716, 2003.

[14] D. H. Kim, M. A. Behlke, S. D. Rose, MI. S. Chang, S. Choi, and J. J. Rossi, "Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy," *Nature Biotechnology*, vol. 23, no. 2, pp. 222–226, 2005.

[15] D. Siolas, C. Lerner, J. Burchard et al., "Synthetic shRNAs as potent RNAi triggers," *Nature Biotechnology*, vol. 23, no. 2, pp. 227–231, 2005.

[16] N. Abe, H. Abe, and Y. Ito, "Dumbbell-shaped nanocircular RNAs for RNA interference," *Journal of the American Chemical Society*, vol. 129, no. 49, pp. 15108–15109, 2007.

[17] J. B. Bramsen, M. B. Laursen, C. K. Damgaard et al., "Improved silencing properties using small internally segmented interfering RNAs," *Nucleic Acids Research*, vol. 35, no. 17, pp. 5886–5897, 2007.

[18] V. A. Korshun, N. B. Pestov, E. V. Nozhevnikova, I. A. Prokhorenko, S. V. Gontarev, and Y. A. Berlin, "Reagents for multiple non-radioactive labelling of oligonucleotides," *Synthetic Communications*, vol. 26, no. 13, pp. 2531–2547, 1996.

[19] M. S. Shchepinov, A. J. Udalova, A. J. Bridgman, and E. M. Southern, "Oligonucleotide dendrimers: synthesis and use as polylabeled DNA probes," *Nucleic Acids Research*, vol. 25, no. 22, pp. 4447–4454, 1997.

[20] M. S. Shchepinov and E. M. Southern, "The synthesis of branched oligonucleotide structures," *Bioorganicheskaya Khimiya*, vol. 24, no. 10, pp. 794–797, 1998.

[21] M. J. Damha and K. K. Ogilvie, "Synthesis and spectroscopic analysis of branched RNA fragments: messenger RNA splicing intermediates," *Journal of Organic Chemistry*, vol. 53, no. 16, pp. 3710–3722, 1988.

[22] E. Utagawa, A. Ohkubo, M. Sekine, and K. Seio, "Synthesis of branched oligonucleotides with three different sequences using an oxidatively removable tritylthio group," *Journal of Organic Chemistry*, vol. 72, no. 22, pp. 8259–8266, 2007.

[23] R. S. Braich and M. J. Damha, "Regiospecific solid-phase synthesis of branched oligonucleotides. Effect of vicinal 2',5'- (or 2',3'-) and 3',5'-phosphodiester linkages on the formation of hairpin DNA," *Bioconjugate Chemistry*, vol. 8, no. 3, pp. 370–377, 1997.

[24] M. J. Damha, K. Ganeshan, R. H.E. Hudson, and S. V. Zabarylo, "Solid-phase synthesis of branched oligoribonucleotides related to messenger RNA splicing intermediates," *Nucleic Acids Research*, vol. 20, no. 24, pp. 6565–6573, 1992.

[25] S. Carriero and M. J. Damha, "Inhibition of pre-mRNA splicing by synthetic branched nucleic acids," *Nucleic Acids Research*, vol. 31, no. 21, pp. 6157–6167, 2003.

[26] M. Grøtli, R. Eritja, and B. Sproat, "Solid-phase synthesis of branched RNA and branched DNA/RNA chimeras," *Tetrahedron*, vol. 53, no. 33, pp. 11317–11346, 1997.

[27] Y. Ueno, M. Takeba, M. Mikawa, and A. Matsuda, "Nucleosides and nucleotides. 182. Synthesis of branched oligodeoxynucleotides with pentaerythritol at the branch point and their thermal stabilization of triplex formation," *Journal of Organic Chemistry*, vol. 64, no. 4, pp. 1211–1217, 1999.

[28] M. D. Sorensen, M. Meldgaard, V. K. Rajwanshi, and J. Wengel, "Branched oligonucleotides containing bicyclic nucleotides as branching points and DNA or LNA as triplex forming branch," *Bioorganic and Medicinal Chemistry Letters*, vol. 10, no. 16, pp. 1853–1856, 2000.

[29] A. Aviñó, M. G. Grimau, M. Frieden, and R. Eritja, "Synthesis and triplex-helix-stabilization properties of branched oligonucleotides carrying 8-aminoadenosine moieties," *Helvetica Chimica Acta*, vol. 87, no. 2, pp. 303–316, 2004.

[30] M. S. Shchepinov, K. U. Mir, J. K. Elder, M. D. Frank-Kamenetskii, and E. M. Southern, "Oligonucleotide dendrimers: stable nano-structures," *Nucleic Acids Research*, vol. 27, no. 15, pp. 3035–3041, 1999.

[31] M. G. Grimau, D. Iacopino, A. Aviñó et al., "Synthesis of branched oligonucleotides as templates for the assembly of nanomaterials," *Helvetica Chimica Acta*, vol. 86, no. 8, pp. 2814–2826, 2003.

[32] S. E. Stanca, A. Ongaro, R. Eritja, and D. Fitzmaurice, "DNA-templated assembly of nanoscale architectures," *Nanotechnology*, vol. 16, no. 9, pp. 1905–1911, 2005.

[33] H. Yang and H. F. Sleiman, "Templated synthesis of highly stable, electroactive, and dynamic metal-DNA branched junctions," *Angewandte Chemie—International Edition*, vol. 47, no. 13, pp. 2443–2446, 2008.

[34] M. Scheffler, A. Dorenbeck, S. Jordan, M. Wüstefeld, and G. Von Kiedrowski, "Self-assembly of trisoligonucleotidyls: the case for nanoacetylene and nano-cyclobutadiene," *Angewandte Chemie—International Edition*, vol. 38, no. 22, pp. 3312–3315, 1999.

[35] T. Horn, C. A. Chang, and M. S. Urdea, "Chemical synthesis and characterization of branched oligodeoxyribonucleotides (bDNA) for use as signal amplifiers in nucleic acid quantification assays," *Nucleic Acids Research*, vol. 25, no. 23, pp. 4842–4849, 1997.

[36] D. R. Sørensen, M. Leirdal, and M. Sioud, "Gene silencing by systemic delivery of synthetic siRNAs in adult mice," *Journal of Molecular Biology*, vol. 327, no. 4, pp. 761–766, 2003.

SAGE-Hindawi Access to Research
Journal of Nucleic Acids
Volume 2011, Article ID 130581, 9 pages
doi:10.4061/2011/130581

Review Article

tRNA/mRNA Mimicry by tmRNA and SmpB in *Trans*-Translation

Daisuke Kurita, Akira Muto, and Hyouta Himeno

Department of Biochemistry and Molecular Biology, Faculty of Agriculture and Life Science, Hirosaki University, Hirosaki 036-8561, Japan

Correspondence should be addressed to Hyouta Himeno, himeno@cc.hirosaki-u.ac.jp

Received 7 October 2010; Accepted 15 December 2010

Academic Editor: Arthur van Aerschot

Since accurate translation from mRNA to protein is critical to survival, cells have developed translational quality control systems. Bacterial ribosomes stalled on truncated mRNA are rescued by a system involving tmRNA and SmpB referred to as *trans*-translation. Here, we review current understanding of the mechanism of *trans*-translation. Based on results obtained by using directed hydroxyl radical probing, we propose a new type of molecular mimicry during *trans*-translation. Besides such chemical approaches, biochemical and cryo-EM studies have revealed the structural and functional aspects of multiple stages of *trans*-translation. These intensive works provide a basis for studying the dynamics of tmRNA/SmpB in the ribosome.

1. Introduction

Translation from the genetic information contained in mRNA to the amino acid sequence of a protein is performed on the ribosome, a large ribonucleoprotein complex composed of three RNA molecules and over 50 proteins. The ribosome is a molecular machine that catalyzes the synthesis of a polypeptide from its substrate, aminoacyl-tRNA. Ribosomes that translate a problematic mRNA, such as that lacking a stop codon, can stall at its 3′ end and produce an incomplete, potentially deleterious protein. *Trans*-translation is known as the highly sophisticated system in bacteria to recycle ribosomes stalled on defective mRNAs and add a short tag-peptide to the C-terminus of the nascent polypeptide as the degradation signal [1–4] (Figure 1). Thus, the tagged polypeptide from truncated mRNA is preferentially degraded by cellular proteases including ClpXP, ClpAP, Lon, FtsH, and Tsp [1, 5–7], and truncated mRNA is released from the stalled ribosomes to be degraded by RNases [8]. The process of *trans*-translation is facilitated by transfer-messenger RNA (tmRNA, also known as 10Sa RNA or SsrA RNA), which is a unique hybrid molecule that functions as both tRNA and mRNA (Figure 2). It comprises two functional domains, the tRNA domain partially mimicking tRNA [9] and the mRNA domain, which includes the coding region for the tag-peptide, surrounded by four pseudoknot structures [10–14]. As predicted from the tRNA-like secondary structure, the 3′ end of tmRNA is aminoacylated by alanyl-tRNA synthetase (AlaRS) like that of canonical tRNA [15, 16]. The function as tRNA is a prerequisite for the function as mRNA, indicating the importance of the elaborate interplay of the two functions [2]. Thus, "*trans*-translation" has been proposed: Ala-tmRNA somehow enters the stalled ribosome, allowing translation to resume by switching the original mRNA to the tag-encoding region on tmRNA. Various questions about the molecular mechanism of this process have been raised. How does tmRNA enter the stalled ribosome in the absence of a codon-anticodon interaction? How is tmRNA switched from the original mRNA in the ribosome? How is the resume codon on tmRNA for the tag-peptide determined? How does tmRNA, 4- or 5-fold larger than tRNA, work in the narrow space in the ribosome?

Several factors, including EF-Tu [17–20], SmpB [21–23], and ribosomal protein S1 [22–24], have been identified as tmRNA-binding proteins. EF-Tu delivers Ala-tmRNA to the ribosome like aminoacyl-tRNA in translation. Unlike S1 [25–27], SmpB serves as an essential factor for *trans*-translation *in vivo* and *in vitro*. It binds to the tRNA-like domain (TLD) of tmRNA [23, 28–30] and ribosome [21] to perform multiple functions during *trans*-translation, including enhancement of aminoacylation efficiency of tmRNA

[22, 23, 31], protection of tmRNA from degradation in the cell [19, 28], and recruitment of tmRNA to the stalled ribosome [21, 23]. NMR studies have revealed that SmpB consists of an antiparallel β-barrel core with three helices and flexible C-terminal tail residues that are disordered in solution [32, 33].

Here, we review recent progress in our understanding of the molecular mechanism of *trans*-translation facilitated by tmRNA and SmpB, which is being revealed by various chemical approaches such as directed hydroxyl radical probing and chemical modification as well as other biochemical and structural studies.

2. *In Vitro Trans*-Translation System

A cell-free *trans*-translation system coupled with poly (U)-dependent polyphenylalanine synthesis was developed using *Escherichia coli* crude cell extracts [2]. Later, several *trans*-translation systems were developed using purified factors from *E. coli* [31, 34, 35] or from *Thermus thermophilus* [25]. These systems have revealed that EF-Tu and SmpB, in addition to the stalled ribosome and Ala-tmRNA, are essential and sufficient for the first few steps of *trans*-translation including the binding of Ala-tmRNA to the ribosome, peptidyl transfer from peptidyl-tRNA to Ala-tmRNA, and decoding of the first codon on tmRNA for the tag peptide. Besides, these systems have also provided a basis for investigating the molecular mechanism of *trans*-translation by chemical approaches.

3. Molecular Mimicries of tRNA and mRNA Revealed by Directed Hydroxyl Radical Probing

Ivanova et al. [36] performed chemical probing to analyze the interaction between SmpB and a ribosome. Bases of rRNA are protected from chemical modification with dimethylsulfate or kethoxal by SmpB, indicating that there are two SmpB-binding sites on the ribosome; one is around the P-site of the small ribosomal subunit and the other is under the L7/L12 stalk of the large ribosomal subunit. The capacity of two SmpB molecules to bind to a ribosome is in agreement with results of other biochemical studies [37, 38]. Gutmann et al. [29] showed a crystal structure of *Aquifex aeolicus* SmpB in complex with the tmRNA fragment corresponding to TLD, which confirmed results of earlier biochemical studies showing that TLD is the crucial binding region of SmpB [23]. It also suggested that SmpB orients toward the decoding center of the small ribosomal subunit and that SmpB structurally mimics the anticodon arm. This is in agreement with a cryo-EM map of the accommodated state complex of ribosome/Ala-tmRNA/SmpB [39–41].

A truncation of the unstructured C-terminal tail of SmpB leads to a loss of *trans*-translation activity [42, 43]. In spite of its functional significance, cryo-EM studies have failed to identify the location of the C-terminal tail of SmpB in the ribosome due to poor resolution. We performed directed hydroxyl radical probing with Fe(II)-BABE to study the sites and modes of binding of *E. coli* SmpB to the ribosome (Figure 3). Fe(II)-BABE is a specific modifier of the cysteine residue of a protein, which generates hydroxyl radicals to cleave the RNA chain. Cleavage sites on RNA can be detected by primer extension, allowing mapping of amino acid residues of a binding protein on an RNA-based macromolecule. This is an excellent chemical approach to study the interaction of a protein with the ribosome [44–47]. We prepared SmpB variants each having a single cysteine residue for attaching it to an Fe(II)-BABE probe. Using directed hydroxyl radical probing, we succeeded in identifying the location of not only the structural domain but also the C-terminal tail of SmpB on the ribosome [48]. It was revealed that there are two SmpB-binding sites in a ribosome, which correspond to the lower halves of the A-site and P-site and that the C-terminal tail of A-site SmpB is aligned along the mRNA path towards the downstream tunnel, while that of P-site SmpB is located almost exclusively around the region of the codon-anticodon interaction in the P-site. This suggests that the C-terminal tail of SmpB mimics mRNA in the A-site and P-site and that these binding sites reflect the pre- and posttranslocation steps of *trans*-translation. The probing signals appear at interval 3, residues of the latter half of the C-terminal tail, suggesting an α helix structure, which has been predicted from the periodical occurrence of positively charged residues [42]. Consequently, the following model has been proposed. The main body of SmpB mimics the lower half of tRNA, and the C-terminal tail of SmpB mimics mRNA both before and after translocation, while the upper half of tRNA is mimicked by TLD. Upon entrance of tmRNA into the stalled ribosome, the C-terminal tail of SmpB may recognize the vacant A-site free of mRNA to trigger *trans* translation. After peptidyl transfer to Ala-tmRNA occurring essentially in the same manner as that in canonical translation, translocation of peptidyl-Ala-tmRNA/SmpB from the A-site to the P-site may occur. During this event, the extended C-terminal tail folds around the region of the codon-anticodon interaction in the P-site, which drives out mRNA from the P-site.

4. Early Stages of *Trans*-Translation

Ala-tmRNA/SmpB forms a complex with EF-Tu and GTP *in vitro*, and this quaternary complex is likely to enter the empty A-site of the stalled ribosome [22]. This complex forms an initial binding complex with the stalled ribosome like the ternary complex of aminoacyl-tRNA, EF-Tu, and GTP does with the translating ribosome. In normal translation, the correct codon-anticodon interaction is recognized by universally conserved 16S rRNA bases, G530, A1492 and A1493, which form the decoding center. When a cognate tRNA binds to the A-site, A1492, and A1493 flip out from the interior of helix 44 of 16S rRNA, and G530 rotates from a syn to an anticonformation to monitor the geometry of the correct codon-anticodon duplex [53]. This induces GTP hydrolysis by EF-Tu, allowing the CCA terminal of tRNA to be accommodated into the peptidyl transferase center. In the context of tRNA mimicry, SmpB should orient toward the decoding center in *trans*-translation. We have recently

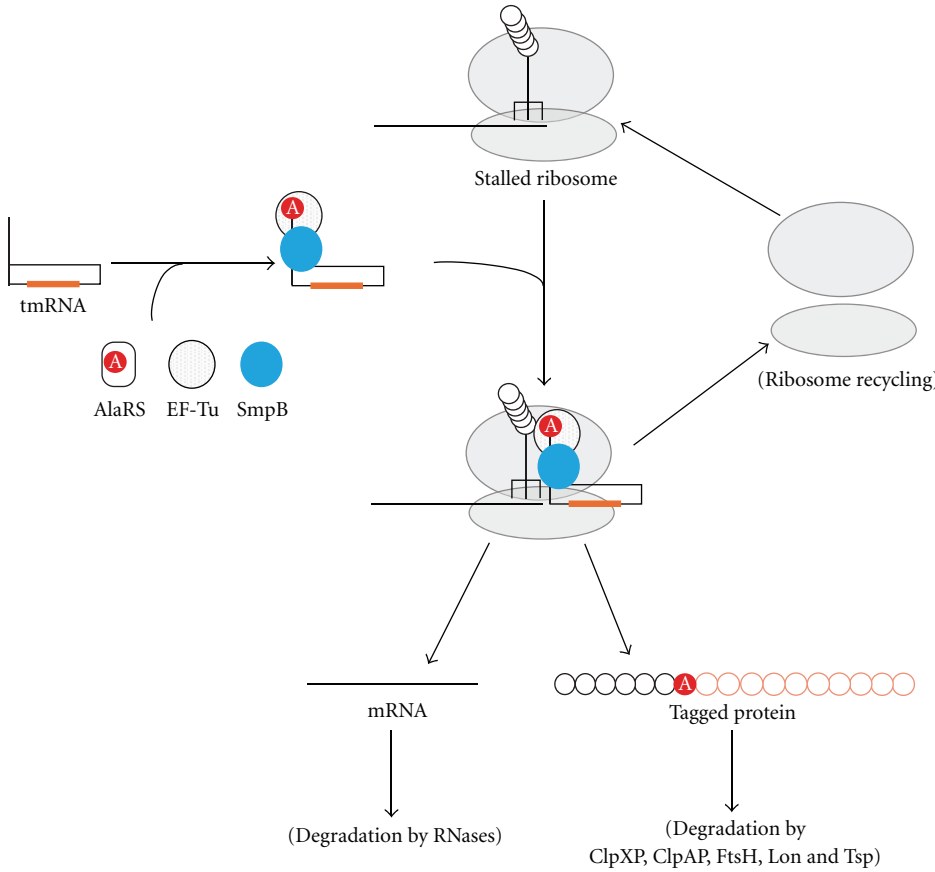


Figure 1: Schematic model of *trans*-translation. Ala-tmRNA receives the nascent polypeptide from peptidyl-tRNA in the P-site of the stalled ribosome to add a tag peptide. *Trans*-translation allows recycling of the ribosomes and promotes degradation of both truncated mRNAs and aberrant polypeptides from truncated mRNAs.

shown that interaction of the C-terminal tail of SmpB with the mRNA path in the ribosome occurs after hydrolysis of GTP by EF-Tu [49]. According to a chemical probing and NMR study, SmpB interacts with G530, A1492, and A1493 [54]. How these bases recognize SmpB to trigger the following GTP hydrolysis is yet to be studied. It should be noted that recent crystal structures have revealed that these bases recognize the A-site ligands (aminoacyl-tRNAs, IF-1, RF-1, RF-2 and RelE) in different ways during translation [50, 55, 56].

Cryo-EM reconstructions of the preaccommodated state of the ribosome/Ala-tmRNA/SmpB/EF-Tu/GDP/ kirromycin complex of *T. thermophilus* have shown that two SmpB molecules present in a complex, one binding to the 50S ribosomal subunit at the GTPase-associated center and the other binding to the 30S subunit near the decoding center [39, 41]. The latter SmpB is not found in the accommodation complex of *T. thermophilus* and *E. coli* [39–41]. Thus, the following model has been proposed: two molecules of SmpB are required for binding of Ala-tmRNA to the stalled ribosome and one of them is released from the ribosome concomitant with the release of EF-Tu after hydrolysis of GTP, so that the 3′-terminal of tmRNA is oriented toward the peptidyl-transferase center. However, several reports have argued against the requirement of two SmpB molecules for *trans*-translation: SmpB has been reported to interact with tmRNA in a 1 : 1 stoichiometry in the cell [57, 58], and crystal structures of SmpB in complex with TLD have been reported to exhibit a 1 : 1 stoichiometry of tmRNA and SmpB [29, 59]. Further studies are required to assess the stoichiometry of SmpB in the preaccommodation state complex.

We have recently shown that the C-terminal tail of SmpB is required for the accommodation of Ala-tmRNA/SmpB into the A-site rather than the initial binding of Ala-tmRNA/SmpB/EF-Tu/GTP to the stalled ribosome [49]. We have also shown that the tryptophan residue at 147 in the middle of the C-terminal tail of *E. coli* SmpB has a crucial role in the step of accommodation. Our results further suggest that the aromatic side chain of Trp147 is required for interaction with rRNA upon accommodation.

It has been shown that *trans*-translation can occur in the middle of an mRNA *in vitro*, although the efficiency of *trans*-translation is dramatically reduced with increase in the length of the 3′ extension from the decoding center [34, 35]. This may be a result of competition of the 3′ extension of mRNA and the C-terminal of A-site SmpB for the mRNA path. The ribosome stalled on the middle of intact mRNA in a cell might be rescued by *trans*-translation via cleavage of

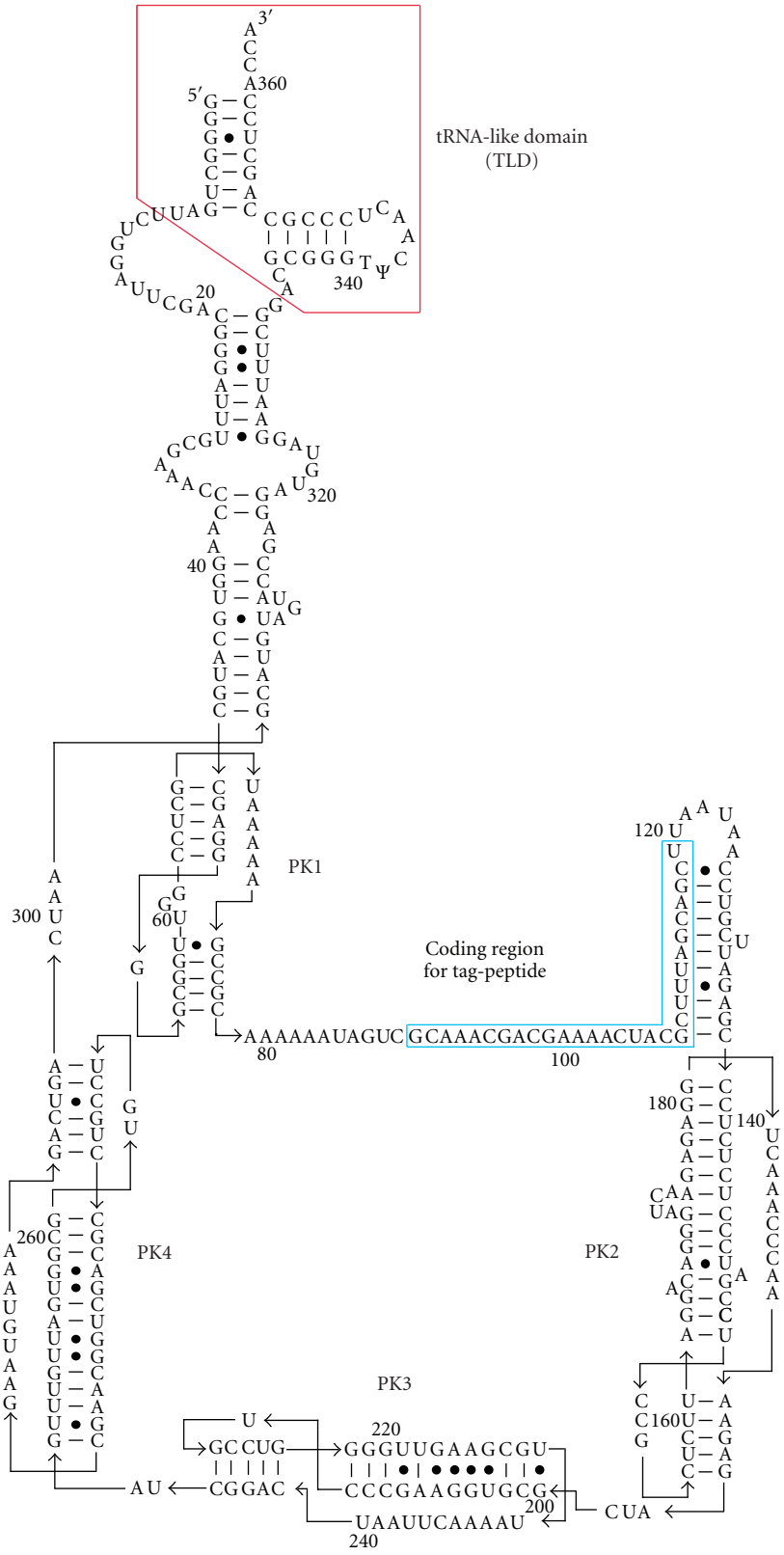


Figure 2: Secondary structure model of tmRNA from *E. coli*. The tRNA-like domain and mRNA domain are highlighted with red and blue, respectively. The tag-encoding sequence is surrounded by four pseudoknot structures (PK1-4).

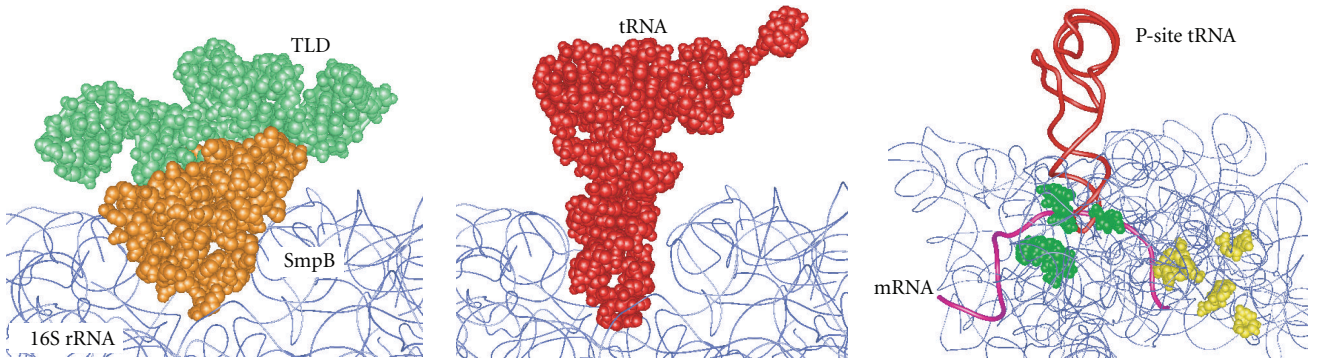


Figure 3: Molecular mimicry of tRNA and mRNA from directed hydroxyl radical probing. (a) Location of TLD-SmpB (left) and tRNA (right) in the ribosomal A-site from directed hydroxyl radical probing [48, 49]. The N-terminal globular domain of SmpB mimics the lower half of tRNA in the A-site. The tertiary structures of TLD-SmpB from *T. thermophilus* [50] and 70S ribosome from *E. coli* [51] were used. (b) Location of the C-terminal tail of SmpB from directed hydroxyl radical probing. Cleavage sites by Fe(II)-tethered A-site and P-site SmpB are colored yellow and green, respectively. The C-terminal tails are located on the mRNA path, suggesting that the C-terminal tail of SmpB mimics mRNA in both the A-site and P-site. P-site SmpB and mRNA are colored red and pink, respectively. The tertiary structure model of 70S ribosome from *T. thermophilus* [52] is used.

mRNA at the A-site [60] or by alternative ribosome rescue systems [61–63].

5. Determination of the Resume Codon

In *trans* translation, the ribosome switches template from a problematic mRNA to tmRNA. How does the stalled ribosome select the first codon on tmRNA without an SD-like sequence? It is reasonable to assume that some structural element on tmRNA is responsible for positioning the resume codon in the decoding center just after translocation of peptidyl-Ala-tmRNA/SmpB from the A-site to the P-site. In *E. coli*, the coding region for the tag peptide starts from position 90 of tmRNA, which is 12 nucleotides downstream of PK1. Indeed, PK1 is important for efficiency of *trans*-translation [14], whereas changing the span between PK1 and the resume codon does not affect determination of the initiation point of tag-translation [64]. A genetic selection experiment has revealed strong base preference in the single-stranded region between PK1 and the resume codon, especially −4 and +1 (position 90) [65]. The importance of this region has also been shown by an *in vitro* study [64]. Several point mutations in this region encompassing −6 to −1 decrease the efficiency of tag-translation, while some of them shift the tag-initiation point by −1 or +1 to a considerable extent [59, 60], indicating that the upstream sequence contains not only the enhancer of *trans*-translation but also the determinant for the tag-initiation point. Evidence for interaction between the upstream region and SmpB has been provided by a study using chemical probing [66]. *E. coli* SmpB protects U at position −5 from chemical modification with CMCT. The structural domain of SmpB rather than the C-terminal tail is involved in this protection. The protection at −5 was suppressed by a point mutation in the TLD critical to SmpB binding, suggesting that SmpB serves to bridge two separate domains of tmRNA to determine the resume codon for tag-translation. Mutations that cause −1 and +1 shifts of the start point of tag-translation also shift the site of protection at −5 from chemical modification by −1 and +1, respectively, indicating the significance of the fixed span between the site of interaction on tmRNA with SmpB and the resume point of translation: translation for the tag-peptide starts from the position 5 nucleotides downstream of the site of interaction with SmpB. Such a functional interaction of the upstream region in tmRNA with SmpB is also supported by the results of another genetic study showing that A-to-C mutation at position 86 of *E. coli* tmRNA that inactivates *trans*-translation both *in vitro* and *in vivo* is suppressed by some double or triple mutations in SmpB [67]. In agreement with these studies, recent cryo-EM studies have suggested that the upstream region in tmRNA interacts with SmpB in the resume (posttranslocation) state [68, 69].

The initiation shift of tag-translation can also be induced by the addition of a 4,5- or 4,6-disubstituted class of aminoglycoside such as paromomycin or neomycin [70, 71], which usually causes miscoding of translation by binding to the decoding center on helix 44 of the small subunit to induce a conformational change in its surroundings [72]. Aminoglycosides also bind at helix 69 of the large subunit, which forms the B2a bridge with helix 44 in close proximity of the decoding center in the small subunit, to inhibit translocation and ribosome recycling by restricting the helical dynamics of helix 69 [73]. Taken together, these findings suggest the significance of interaction of the proximity of the decoding center with any portion of SmpB or tmRNA for precise tag-translation. It should be noted that hygromycin B, which binds only to helix 44, does not induce initiation shift of tag-translation [71].

6. Trajectories of tmRNA/SmpB

Along with the functional mimicry of TLD/SmpB, a similar behavior of tmRNA/SmpB to that of canonical tRNA+mRNA in the ribosome through several hybrid states,

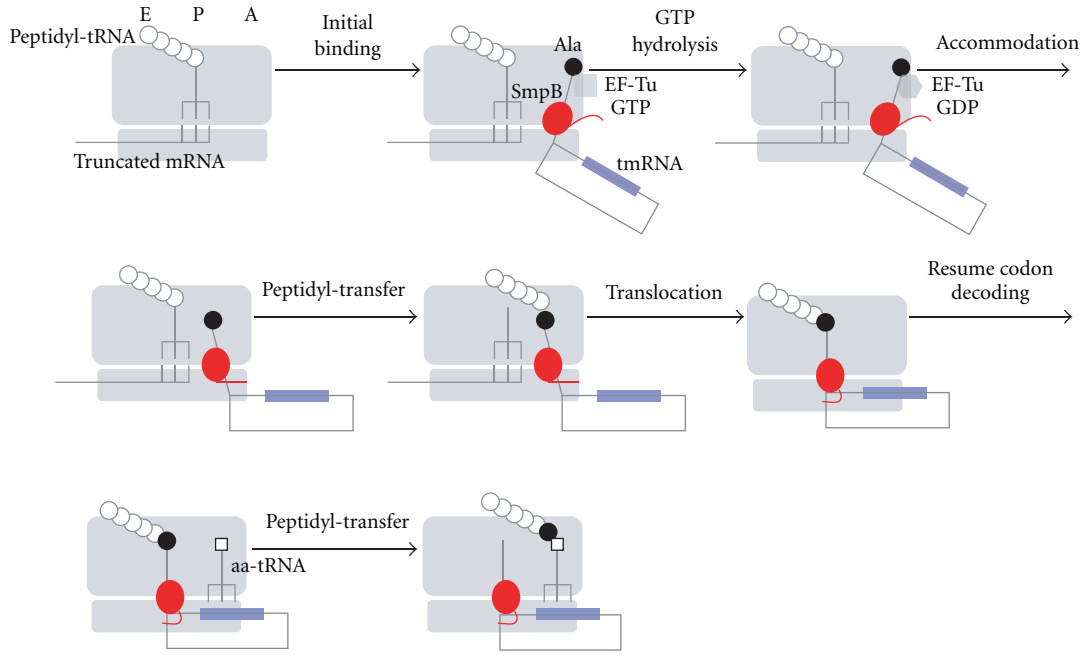


Figure 4: A model of the early stages of *trans*-translation. The C-terminal tail of SmpB is not located on the mRNA path in the processes before accommodation. After GTP hydrolysis by EF-Tu, the C-terminal tail is located on the mRNA path mimicking mRNA to recognize the stalled ribosome free of mRNA. Following translocation of tmRNA/SmpB from the A-site to P-site, the C-terminal tail undergoes drastic conformational change to accommodate the resume codon of tmRNA into the decoding center. SmpB and the tag-encoding region are shown by red and blue, respectively. White circles indicate amino acids encoded by truncated mRNA, and a white square indicates amino acid designated by the resume codon of tmRNA.

A/T, A/A, A/P, P/P, and P/E, has been assumed. Cryo-EM studies have shown the location of the complex of tmRNA with the main body of SmpB in the A/T and A/A states [39, 40], and a directed hydroxyl radical probing has revealed the positions of SmpB in the A/A and P/P states [48]. The existence of stable SmpB binding sites in the A-site and P-site suggests the requirement of translocation, as in canonical translation. It might possibly involve EF-G. Concomitantly with translocation, mRNA and P-site tRNA are released from the stalled ribosome [74]. Considering the different C-terminal tail structures of A-site SmpB and P-site SmpB, the C-terminal tail would somehow undergo conformational change from the extended form to the folded form [48].

The next translocation is thought to move tmRNA/SmpB to the E-site. These ribosomal processes should involve extensive changes in the conformation of tmRNA [75] as well as in the modes of interactions of tmRNA with SmpB and the ribosome [76, 77]. According to chemical probing studies, secondary structure elements of tmRNA remain intact in a few steps of *trans*-translation including pre- and posttranslocation states [77–79]. Another study has suggested 1 : 1 stoichiometry of tmRNA to SmpB throughout the processes of translation for the tag peptide [80]. Recently, the movement of tRNA during translocation has been revealed by using time-resolved cryo-EM [81]. Not only classic and hybrid states but also various novel intermediate states of tRNAs were revealed. Although the intermediate states during *trans*-translation remain unclear, results of future structural studies including chemical approaches should reveal tmRNA/SmpB and ribosome dynamics.

7. Conclusion

Various chemical approaches in addition to cryo-EM and X-ray crystallographic studies have been revealing the molecular mechanism of *trans*-translation. tmRNA forms a ribonucleoprotein complex with SmpB, which plays an essential role in *trans*-translation. Based on a directed hydroxyl radical probing towards SmpB, we have proposed a novel molecular mechanism of *trans*-translation (Figure 4). In this model, an elegant collaboration of a hybrid RNA molecule of tRNA and mRNA and a protein mimicking a set of tRNA and mRNA facilitates *trans*-translation. Initially, a quaternary complex of Ala-tmRNA, SmpB, EF-Tu, and GTP may enter the vacant A-site of the stalled ribosome to trigger *trans*-translation, when a set of Ala-TLD of tmRNA and the main body of SmpB mimicking the upper and lower halves of aminoacyl-tRNA, respectively, recognizes the A-site free of tRNA. After hydrolysis of GTP by EF-Tu, the C-terminal tail of SmpB mimicking mRNA interacts with the decoding center and the downstream mRNA path free of mRNA, allowing Ala-TLD/SmpB to be accommodated. While several proteins including SmpB have been proposed to mimic tRNA or its portion, SmpB is the first protein that has been shown to mimic mRNA. SmpB is also the first protein of which stepwise movements in the ribosome are assumed to mimic those of tRNA in the translating ribosome.

Our model depicts an outline of the *trans*-translation processes in the ribosome, although the following issues should be addressed. How do the intermolecular interactions between tmRNA and ribosome, between tmRNA and SmpB, and between ribosome and SmpB as well as the intramolecular interactions within tmRNA and within SmpB change during the course of the *trans*-translation processes? Is EF-G required for translocation of tmRNA/SmpB having neither an anticodon nor the corresponding codon from the A-site to the P-site? If EF-G is required, how does it promote translocation? These questions remain to be answered in the future works.

Abbreviations

AlaRS: Alanyl-tRNA synthetase
BABE: 1-(*p*-bromoacetamidobenzyl)-ethylenediaminetetraacetic acid
CMCT: 1-cyclohexyl-3-(2-morpholinoethyl) carbodiimide metho-*p*-toluene sulfonate
cryo-EM: Cryo-electron microscopy
EF-Tu: Elongation factor Tu
EF-G: Elongation factor G
PK: Pseudoknot
TLD: tRNA-like domain.

Acknowledgments

The authors are grateful to all those who have been involved in this work. This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture, Japan to H. Himeno, Grants-in-Aid for Scientific Research (B) and (C) from the Japan Society for the Promotion of Science, Grant for Hirosaki University Institutional Research to H. Himeno and A. Muto, and Grants-in-Aid for Young Scientist (Start-up) to D. Kurita.

References

[1] K. C. Keiler, P. R. H. Waller, and R. T. Sauer, "Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA," *Science*, vol. 271, no. 5251, pp. 990–993, 1996.

[2] H. Himeno, M. Sato, T. Tadaki, M. Fukushima, C. Ushida, and A. Muto, "In vitro *trans* translation mediated by alanine-charged 10Sa RNA," *Journal of Molecular Biology*, vol. 268, no. 5, pp. 803–808, 1997.

[3] A. Muto, M. Sato, T. Tadaki, M. Fukushima, C. Ushida, and H. Himeno, "Structure and function of 10Sa RNA: *trans*-translation system," *Biochimie*, vol. 78, no. 11-12, pp. 985–991, 1996.

[4] A. Muto, C. Ushida, and H. Himeno, "A bacterial RNA that functions as both a tRNA and an mRNA," *Trends in Biochemical Sciences*, vol. 23, no. 1, pp. 25–29, 1998.

[5] S. Gottesman, E. Roche, Y. Zhou, and R. T. Sauer, "The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system," *Genes and Development*, vol. 12, no. 9, pp. 1338–1347, 1998.

[6] C. Herman, D. Thévenet, P. Bouloc, G. C. Walker, and R. D'Ari, "Degradation of carboxy-terminal-tagged cytoplasmic proteins by the *Escherichia coli* protease HflB (FtsH)," *Genes and Development*, vol. 12, no. 9, pp. 1348–1355, 1998.

[7] J. S. Choy, L. L. Aung, and A. W. Karzai, "Lon protease degrades transfer-messenger RNA-tagged proteins," *Journal of Bacteriology*, vol. 189, no. 18, pp. 6564–6571, 2007.

[8] Y. Yamamoto, T. Sunohara, K. Jojima, T. Inada, and H. Aiba, "SsrA-mediated *trans*-translation plays a role in mRNA quality control by facilitating degradation of truncated mRNAs," *RNA*, vol. 9, no. 4, pp. 408–418, 2003.

[9] B. Felden, K. Hanawa, J. F. Atkins et al., "Presence and location of modified nucleotides in *Escherichia coli* tmRNA: structural mimicry with tRNA acceptor branches," *EMBO Journal*, vol. 17, no. 11, pp. 3188–3196, 1998.

[10] K. P. Williams and D. P. Bartel, "Phylogenetic analysis of tmRNA secondary structure," *RNA*, vol. 2, no. 12, pp. 1306–1310, 1996.

[11] B. Felden, H. Himeno, A. Muto, J. F. Atkins, and R. F. Gesteland, "Structural organization of *Escherichia coli* tmRNA," *Biochimie*, vol. 78, no. 11-12, pp. 979–983, 1996.

[12] B. Felden, H. Himeno, A. Muto, J. P. McCutcheon, J. F. Atkins, and R. F. Gesteland, "Probing the structure of the *Escherichia coli* 10Sa RNA (tmRNA)," *RNA*, vol. 3, no. 1, pp. 89–103, 1997.

[13] R. P. Hickerson, C. D. Watkins-Sims, C. J. Burrows, J. F. Atkins, R. F. Gesteland, and B. Felden, "A nickel complex cleaves uridine in folded RNA structures: application to *E. coli* tmRNA and related engineered molecules," *Journal of Molecular Biology*, vol. 279, no. 3, pp. 577–587, 1998.

[14] N. Nameki, B. Felden, J. F. Atkins, R. F. Gesteland, H. Himeno, and A. Muto, "Functional and structural analysis of a pseudoknot upstream of the tag-encoded sequence in *E. coli* tmRNA," *Journal of Molecular Biology*, vol. 286, no. 3, pp. 733–744, 1999.

[15] C. Ushida, H. Himeno, T. Watanabe, and A. Muto, "TRNA-like structures in 10Sa RNAs of *Mycoplasma capricolum* and *Bacillus subtilis*," *Nucleic Acids Research*, vol. 22, no. 16, pp. 3392–3396, 1994.

[16] Y. Komine, M. Kitabatake, T. Yokogawa, K. Nishikawa, and H. Inokuchi, "A tRNA-like structure is present in 10Sa RNA, a small stable RNA from *Escherichia coli*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 91, no. 20, pp. 9223–9227, 1994.

[17] J. Rudinger-Thirion, R. Giege, and B. Felden, "Aminoacylated tmRNA from *Escherichia coli* interacts with prokaryotic elongation factor Tu," *RNA*, vol. 5, no. 8, pp. 989–992, 1999.

[18] S. Barends, J. Wower, and B. Kraal, "Kinetic parameters for tmRNA binding to alanyl-tRNA synthetase and elongation factor Tu from *Escherichia coli*," *Biochemistry*, vol. 39, no. 10, pp. 2652–2658, 2000.

[19] K. Hanawa-Suetsugu, V. Bordeau, H. Himeno, A. Muto, and B. Felden, "Importance of the conserved nucleotides around the tRNA-like structure of *Escherichia coli* transfer-messenger RNA for protein tagging," *Nucleic Acids Research*, vol. 29, no. 22, pp. 4663–4673, 2001.

[20] V. G. Stepanov and J. Nyborg, "tmRNA from *Thermus thermophilus* interaction with alanyl-tRNA synthetase and elongation factor Tu," *European Journal of Biochemistry*, vol. 270, no. 3, pp. 463–475, 2003.

[21] A. W. Karzai, M. M. Susskind, and R. T. Sauer, "SmpB, a unique RNA-binding protein essential for the peptide-tagging activity of SsrA (tmRNA)," *EMBO Journal*, vol. 18, no. 13, pp. 3793–3799, 1999.

[22] S. Barends, A. W. Karzai, R. T. Sauer, J. Wower, and B. Kraal, "Simultaneous and functional binding of SmpB and

EF-Tu·GTP to the alanyl acceptor arm of tmRNA," *Journal of Molecular Biology*, vol. 314, no. 1, pp. 9–21, 2001.

[23] K. Hanawa-Suetsugu, M. Takagi, H. Inokuchi, H. Himeno, and A. Muto, "SmpB functions in various steps of *trans*-translation," *Nucleic Acids Research*, vol. 30, no. 7, pp. 1620–1629, 2002.

[24] I. K. Wower, C. W. Zwieb, S. A. Guven, and J. Wower, "Binding and cross-linking of tmRNA to ribosomal protein S1, on and off the *Escherichia coli* ribosome," *EMBO Journal*, vol. 19, no. 23, pp. 6612–6621, 2000.

[25] K. Takada, C. Takemoto, M. Kawazoe et al., "In vitro *trans*-translation of *Thermus thermophilus*: ribosomal protein S1 is not required for the early stage of *trans*-translation," *RNA*, vol. 13, no. 4, pp. 503–510, 2007.

[26] H. Qi, Y. Shimizu, and T. Ueda, "Ribosomal Protein S1 Is not Essential for the *trans*-translation Machinery," *Journal of Molecular Biology*, vol. 368, no. 3, pp. 845–852, 2007.

[27] M. Saguy, R. Gillet, P. Skorski, S. Hermann-Le Denmat, and B. Felden, "Ribosomal protein S1 influences *trans*-translation in vitro and in vivo," *Nucleic Acids Research*, vol. 35, no. 7, pp. 2368–2376, 2007.

[28] J. Wower, C. W. Zwieb, D. W. Hoffman, and I. K. Wower, "SmpB: a protein that binds to double-stranded segments in tmRNA and tRNA," *Biochemistry*, vol. 41, no. 28, pp. 8826–8836, 2002.

[29] S. Gutmann, P. W. Haebel, L. Metzinger, M. Sutter, B. Felden, and N. Ban, "Crystal structure of the transfer-RNA domain of transfer-messenger RNA in complex with SmpB," *Nature*, vol. 424, no. 6949, pp. 699–703, 2003.

[30] N. Nameki, T. Someya, S. Okano et al., "Interaction analysis between tmRNA and SmpB from *Thermus thermophilus*," *Journal of Biochemistry*, vol. 138, no. 6, pp. 729–739, 2005.

[31] Y. Shimizu and T. Ueda, "The role of SmpB protein in *trans*-translation," *FEBS Letters*, vol. 514, no. 1, pp. 74–77, 2002.

[32] G. Dong, J. Nowakowski, and D. W. Hoffman, "Structure of small protein B: the protein component of the tmRNA-SmpB system for ribosome rescue," *EMBO Journal*, vol. 21, no. 7, pp. 1845–1854, 2002.

[33] T. Someya, N. Nameki, H. Hosoi et al., "Solution structure of a tmRNA-binding protein, SmpB, from *Thermus thermophilus*," *FEBS Letters*, vol. 535, no. 1–3, pp. 94–100, 2003.

[34] N. Ivanova, M. Y. Pavlov, B. Felden, and M. Ehrenberg, "Ribosome rescue by tmRNA requires truncated mRNAs," *Journal of Molecular Biology*, vol. 338, no. 1, pp. 33–41, 2004.

[35] K. Asano, D. Kurita, K. Takada, T. Konno, A. Muto, and H. Himeno, "Competition between *trans*-translation and termination or elongation of translation," *Nucleic Acids Research*, vol. 33, no. 17, pp. 5544–5552, 2005.

[36] N. Ivanova, M. Y. Pavlov, E. Bouakaz, M. Ehrenberg, and L. H. Schiavone, "Mapping the interaction of SmpB with ribosomes by footprinting of ribosomal RNA," *Nucleic Acids Research*, vol. 33, no. 11, pp. 3529–3539, 2005.

[37] M. Hallier, N. Ivanova, A. Rametti, M. Pavlov, M. Ehrenberg, and B. Felden, "Pre-binding of small protein B to a stalled ribosome triggers *trans*-translation," *Journal of Biological Chemistry*, vol. 279, no. 25, pp. 25978–25985, 2004.

[38] M. Hallier, J. Desreac, and B. Felden, "Small protein B interacts with the large and the small subunits of a stalled ribosome during *trans*-translation," *Nucleic Acids Research*, vol. 34, no. 6, pp. 1935–1943, 2006.

[39] S. Kaur, R. Gillet, W. Li, R. Gursky, and J. Frank, "Cryo-EM visualization of transfer messenger RNA with two SmpBs in a stalled ribosome," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 44, pp. 16484–16489, 2006.

[40] K. Cheng, N. Ivanova, S. H. W. Scheres et al., "tmRNA·SmpB complex mimics native aminoacyl-tRNAs in the A site of stalled ribosomes," *Journal of Structural Biology*, vol. 169, no. 3, pp. 342–348, 2010.

[41] F. Weis, P. Bron, J. P. Rolland, D. Thomas, B. Felden, and R. Gillet, "Accommodation of tmRNA-SmpB into stalled ribosomes: a cryo-EM study," *RNA*, vol. 16, no. 2, pp. 299–306, 2010.

[42] Y. Jacob, S. M. Sharkady, K. Bhardwaj, A. Sanda, and K. P. Williams, "Function of the SmpB tail in transfer-messenger RNA translation revealed by a nucleus-encoded form," *Journal of Biological Chemistry*, vol. 280, no. 7, pp. 5503–5509, 2005.

[43] T. R. Sundermeier, D. P. Dulebohn, H. J. Cho, and A. W. Karzai, "A previously uncharacterized role for small protein B (SmpB) in transfer messenger RNA-mediated *trans*-translation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 7, pp. 2316–2321, 2005.

[44] K. S. Wilson and H. F. Noller, "Mapping the position of translational elongation factor EF-G in the ribosome by directed hydroxyl radical probing," *Cell*, vol. 92, no. 1, pp. 131–139, 1998.

[45] K. S. Wilson, K. Ito, H. F. Noller, and Y. Nakamura, "Functional sites of interaction between release factor RF1 and the ribosome," *Nature Structural Biology*, vol. 7, no. 10, pp. 866–870, 2000.

[46] A. Dallas and H. F. Noller, "Interaction of translation initiation factor 3 with the 30S ribosomal subunit," *Molecular Cell*, vol. 8, no. 4, pp. 855–864, 2001.

[47] L. Lancaster, M. C. Kiel, A. Kaji, and H. F. Noller, "Orientation of ribosome recycling factor in the ribosome from directed hydroxyl radical probing," *Cell*, vol. 111, no. 1, pp. 129–140, 2002.

[48] D. Kurita, R. Sasaki, A. Muto, and H. Himeno, "Interaction of SmpB with ribosome from directed hydroxyl radical probing," *Nucleic Acids Research*, vol. 35, no. 21, pp. 7248–7255, 2007.

[49] D. Kurita, A. Muto, and H. Himeno, "Role of the C-terminal tail of SmpB in the early stage of *trans*-translation," *RNA*, vol. 16, no. 5, pp. 980–990, 2010.

[50] A. Weixlbaumer, H. Jin, C. Neubauer et al., "Insights into translational termination from the structure of RF2 bound to the ribosome," *Science*, vol. 322, no. 5903, pp. 953–956, 2008.

[51] A. Vila-Sanjurjo, W. K. Ridgeway, V. Seymaner et al., "X-ray crystal structures of the WT and a hyper-accurate ribosome from *Escherichia coli*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 15, pp. 8682–8687, 2003.

[52] G. Yusupova, L. Jenner, B. Rees, D. Moras, and M. Yusupov, "Structural basis for messenger RNA movement on the ribosome," *Nature*, vol. 444, no. 7117, pp. 391–394, 2006.

[53] J. M. Ogle, D. E. Brodersen, W. M. Clemons Jr., M. J. Tarry, A. P. Carter, and V. Ramakrishnan, "Recognition of cognate transfer RNA by the 30S ribosomal subunit," *Science*, vol. 292, no. 5518, pp. 897–902, 2001.

[54] S. Nonin-Lecomte, N. Germain-Amiot, R. Gillet et al., "Ribosome hijacking: a role for small protein B during *trans*-translation," *EMBO Reports*, vol. 10, no. 2, pp. 160–165, 2009.

[55] A. Korostelev, D. N. Ermolenko, and H. F. Noller, "Structural dynamics of the ribosome," *Current Opinion in Chemical Biology*, vol. 12, no. 6, pp. 674–683, 2008.

[56] C. Neubauer, Y. G. Gao, K. R. Andersen et al., "The structural basis for mRNA recognition and cleavage by the ribosome-dependent endonuclease RelE," *Cell*, vol. 139, no. 6, pp. 1084–1095, 2009.

[57] S. D. Moore and R. T. Sauer, "Ribosome rescue: tmRNA tagging activity and capacity in *Escherichia coli*," *Molecular Microbiology*, vol. 58, no. 2, pp. 456–466, 2005.

[58] T. R. Sundermeier and A. W. Karzai, "Functional SmpB-ribosome interactions require tmRNA," *Journal of Biological Chemistry*, vol. 282, no. 48, pp. 34779–34786, 2007.

[59] Y. Bessho, R. Shibata, S. I. Sekine et al., "Structural basis for functional mimicry of long-variable-arm tRNA by transfer-messenger RNA," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 20, pp. 8293–8298, 2007.

[60] K. Pedersen, A. V. Zavialov, M. YU. Pavlov, J. Elf, K. Gerdes, and M. Ehrenberg, "The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site," *Cell*, vol. 112, no. 1, pp. 131–140, 2003.

[61] H. Himeno, "Novel factor rescues ribosomes trapped on non-stop mRNAs," *Molecular Microbiology*, vol. 78, no. 4, pp. 789–791, 2010.

[62] Y. Chadani, K. Ono, S. -I. Ozawa et al., "Ribosome rescue by *Escherichia coli* ArfA (YhdL) in the absence of *trans*-translation system," *Molecular Microbiology*, vol. 78, no. 4, pp. 796–808, 2010.

[63] Y. Handa, N. Inaho, and N. Nameki, "YaeJ is a novel ribosome-associated protein in *Escherichia coli* that can hydrolyze peptidyl-tRNA on stalled ribosomes," *Nucleic Acids Research*. In press.

[64] S. Lee, M. Ishii, T. Tadaki, A. Muto, and H. Himeno, "Determinants on tmRNA for initiating efficient and precise *trans*-translation: some mutations upstream of the tag-encoding sequence of *Escherichia coli* tmRNA shift the initiation point of *trans*-translation in vitro," *RNA*, vol. 7, no. 7, pp. 999–1012, 2001.

[65] K. P. Williams, K. A. Martindale, and D. P. Bartel, "Resuming translation on tmRNA: a unique mode of determining a reading frame," *EMBO Journal*, vol. 18, no. 19, pp. 5423–5433, 1999.

[66] T. Konno, D. Kurita, K. Takada, A. Muto, and H. Himeno, "A functional interaction of SmpB with tmRNA for determination of the resuming point of *trans*-translation," *RNA*, vol. 13, no. 10, pp. 1723–1731, 2007.

[67] T. Watts, D. Cazier, D. Healey, and A. Buskirk, "SmpB contributes to reading frame selection in the translation of transfer-messenger RNA," *Journal of Molecular Biology*, vol. 391, no. 2, pp. 275–281, 2009.

[68] J. Fu, Y. Hashem, I. Wower et al., "Visualizing the transfer-messenger RNA as the ribosome resumes translation," *EMBO Journal*, vol. 29, no. 22, pp. 3819–3825, 2010.

[69] F. Weis, P. Bron, E. Giudice et al., "TmRNA-SmpB: a journey to the centre of the bacterial ribosome," *EMBO Journal*, vol. 29, no. 22, pp. 3810–3818, 2010.

[70] T. Takahashi, T. Konno, A. Muto, and H. Himeno, "Various effects of paromomycin on tmRNA-directed *trans*-translation," *Journal of Biological Chemistry*, vol. 278, no. 30, pp. 27672–27680, 2003.

[71] T. Konno, T. Takahashi, D. Kurita, A. Muto, and H. Himeno, "A minimum structure of aminoglycosides that causes an initiation shift of *trans*-translation," *Nucleic Acids Research*, vol. 32, no. 14, pp. 4119–4126, 2004.

[72] A. P. Carter, W. M. Clemons, D. E. Brodersen, R. J. Morgan-Warren, B. T. Wimberly, and V. Ramakrishnan, "Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics," *Nature*, vol. 407, no. 6802, pp. 340–348, 2000.

[73] M. A. Borovinskaya, R. D. Pai, W. Zhang et al., "Structural basis for aminoglycoside inhibition of bacterial ribosome recycling," *Nature Structural and Molecular Biology*, vol. 14, no. 8, pp. 727–732, 2007.

[74] N. Ivanova, M. Y. Pavlov, and M. Ehrenberg, "tmRNA-induced release of messenger RNA from stalled ribosomes," *Journal of Molecular Biology*, vol. 350, no. 5, pp. 897–905, 2005.

[75] I. K. Wower, C. Zwieb, and J. Wower, "Transfer-messenger RNA unfolds as it transits the ribosome," *RNA*, vol. 11, no. 5, pp. 668–673, 2005.

[76] O. V. Shpanchenko, M. I. Zvereva, P. V. Ivanov et al., "Stepping transfer messenger RNA through the ribosome," *Journal of Biological Chemistry*, vol. 280, no. 18, pp. 18368–18374, 2005.

[77] N. Ivanova, M. Lindell, M. Pavlov, L. H. Schiavone, E. G. H. Wagner, and M. Ehrenberg, "Structure probing of tmRNA in distinct stages of *trans*-translation," *RNA*, vol. 13, no. 5, pp. 713–722, 2007.

[78] E. Y. Bugaeva, S. Surkov, A. V. Golovin et al., "Structural features of the tmRNA-ribosome interaction," *RNA*, vol. 15, no. 12, pp. 2312–2320, 2009.

[79] O. V. Shpanchenko, A. V. Golovin, E. Y. Bugaeva, L. A. Isaksson, and O. A. Dontsova, "Structural aspects of *trans*-translation," *IUBMB Life*, vol. 62, no. 2, pp. 120–124, 2010.

[80] E. Y. Bugaeva, O. V. Shpanchenko, B. Felden, L. A. Isaksson, and O. A. Dontsova, "One SmpB molecule accompanies tmRNA during its passage through the ribosomes," *FEBS Letters*, vol. 582, no. 10, pp. 1532–1536, 2008.

[81] N. Fischer, A. L. Konevega, W. Wintermeyer, M. V. Rodnina, and H. Stark, "Ribosome dynamics and tRNA movement by time-resolved electron cryomicroscopy," *Nature*, vol. 466, no. 7304, pp. 329–333, 2010.