Functional Nucleic Acids for Analytical Applications

Yingfu Li • Yi Lu
Editors

Functional Nucleic Acids for Analytical Applications

Springer

Editors
Yingfu Li
Department of Biochemistry and
Biomedical Sciences
McMaster University
Hamilton, Ontario
Canada
liying@mcmaster.ca

Yi Lu
Department of Chemistry
University of Illinois at Urbana-Champaign
Urbana, Illinois
USA
yi-lu@illinois.edu

Cover illustration: Janat Sinn-Hanlon, Imaging Technology Group, Beckman Institute, University of Illinois at Urbana-Champaign.

ISBN: 978-0-387-73710-2 e-ISBN: 978-0-387-73711-9
DOI: 10.1007/978-0-387-73711-9

Library of Congress Control Number: 2008940063

Printed on acid-free paper

springer.com

Preface

Nature long ago solved the problem of finding sequences that code for useful structures, through the endless iteration of the simple algorithm at the heart of Darwinian evolution: variation, selection, reproduction. The deliberate application of this algorithm to the laboratory evolution of useful molecules is a recent development, but the power of this approach is already evident. My introduction to this field came almost 20 years ago when I wanted to explore the rich biochemistry implicit in the strong version of the RNA world hypothesis. Attempting to evolve RNAs that could carry out the key functions of RNA-based life seemed ambitious, so Andrew Ellington and I decided to start with the simple project of selecting for sequences capable of recognizing a given target. In 1990 we showed that it was indeed possible to evolve new RNAs that could bind to and distinguish between closely related small molecules. This was incredibly exciting to us because it was clear that the selected RNAs had to fold into defined three-dimensional structures that contained highly specific ligand-binding sites. Moreover, these ligand-binding RNAs, which we called aptamers, had been selected from a small (only 10^{15}) sample of completely random RNA sequences, implying that functional RNAs were relatively common in sequence space, and that some day really useful aptamers might be evolved! In parallel with our work, Craig Tuerk and Larry Gold found that unexpected sequence variants of a stem-loop RNA had emerged from a randomized population selected for binding to an RNA-binding phage coat protein, and their subsequent work showed that RNAs could be evolved that would bind to almost any protein target. These early findings were soon followed by the evolution of DNA aptamers, novel ribozymes and DNAzymes, and allosterically controlled ribozymes ("aptazymes"). Since then there has been an explosion of work devoted to the evolution of increasingly sophisticated and useful aptamers and nucleic acid catalysts, collectively referred to as functional nucleic acids or FNAs. Perhaps the most scientifically interesting and surprising application of FNAs has been the exciting effort aimed at the development of biosensors and other analytical applications, such as bioseparations, signal amplification, and signal processing, and it is this work that is summarized in a series of thorough and insightful reviews in the present volume.

The book begins with three excellent reviews, the first (by the editors of this volume) introducing the analytical applications of FNAs that are discussed in detail in Parts II and III of the book, the second covering natural riboswitches (Nature's own

RNA-based biosensors) and ribozymes, and the third covering artificially evolved aptamers, ribozymes, and DNAzymes. In Part II we see the remarkable array of amplification and detection technologies that have been coupled to FNAs to allow accurate and sensitive detection of a vast range of target analytes. These methods include a variety of fluorescence and colorimetric and other optical methods, as well as electrochemistry and catalytic signal amplification. Part III on emerging analytical applications is the most forward-looking part of the book, covering diverse topics ranging from aptamer-based separations, to massively parallel microarray detectors, to computational devices and nanomachines built from functional nucleic acids. The myriad of clever ways in which FNAs are now being used is truly remarkable, and this volume provides a wonderful overview for the reader interested in the current state of the art in this rapidly developing field.

Massachusetts General Hospital, USA Jack W. Szostak

Contents

Part I
Overview of Functional Nucleic Acids and Their Analytical Applications

Chapter 1
Introductory Remarks

Yi Lu and Yingfu Li

Abstract The emergence of a large number of natural and artificial functional nucleic acids (FNAs; aptamers and nucleic acid enzymes, collectively termed functional nucleic acids in this book) has generated tremendous enthusiasm and new opportunities for molecular scientists from diverse disciplines to devise new concepts and applications. In this volume, we have assembled some leading experts to provide a timely account of recent progress in sensing and other analytical applications that explore functional nucleic acids.

It is widely known that nucleic acids are the blueprint of life and the foundation of modern biology: they are the hereditary material for the storage and transmission of genetic information in all living organisms on Earth. In some circles, nucleic acids are also known to have the ability to perform other interesting functions, including catalysis and ligand binding. The first discovery in this specific arena was made in early 1980s when some natural RNA molecules were shown to function as enzymes (ribozymes).[1,2] Several years later, several pioneering researchers came up with the idea to perform selection experiments with DNA or RNA in test tubes using a combinatorial technique known as "in vitro selection" or "SELEX" (systematic evolution of ligands by exponential enrichment).[3–5] Subsequent efforts in applying this elegant technique to large random-sequence DNA or RNA pools have led to the creation of a great number of man-made ribozymes, deoxyribozymes (DNAzymes), and aptamers (molecular receptors made of DNA or RNA).[6–12] More recently, many natural RNA molecules, known as "riboswitches," have joined the rank of functional nucleic acids (FNAs). These tiny "molecular wonders" have the ability to regulate gene expression through binding to some important metabolites such as

Y. Lu
Department of Chemistry, University of Illinois at Urbana-Champaign, USA
yi-lu@illinois.edu

Y. Li
Department of Biochemistry and Biomedical Sciences,
McMaster University, Hamilton, Ontario, Canada
liying@mcmaster.ca

Y. Li and Y. Lu (eds.), *Functional Nucleic Acids for Analytical Applications*,

amino acids and nucleotsides.[13–16] We think it is fair to say that the discovery of new functions of nucleic acids has been one of the most fascinating topics in science over the past 25 years.

The existence of many ribozymes and riboswitches in Nature (see Chapter 2, by Lafontaine and coworkers, for a review on natural ribozymes and riboswitches) and our ability to perform test tube evolution experiments to create man-made enzymes and receptors from DNA or RNA (see Chapter 3, by Scott Silverman, for a review on artificial nucleic acid enzymes and aptamers), have generated tremendous enthusiasm in the scientific community. More and more researchers have become interested in studying the fundamental properties of FNAs and examining them for many novel applications in areas including therapeutics,[17–20] molecular imaging,[21] drug screening,[22–24] affinity separation,[25] materials science,[26] nanotechnology,[27–29] and biosensing.[30–36] A timely summary of accomplishments and critical outlook for the future in these areas will be valuable to a wide range of readers who are, or may become, interested in FNAs. However, it is impossible for one book to cover all aspects of FNA-based applications. In this book, we chose to cover FNAs in sensing and other analytical applications because we have witnessed explosive growth of research activities in this area (Fig. 1.1).

To provide a book with the most useful information and authoritative reviews, we have assembled some leading FNA experts whose research activities have contributed significantly to the creation and expansion of each topic covered. The book covers many aspects of functional nucleic acids, from their structures and creation (Chapters 2 and 3) all the way to their applications in analytical chemistry and beyond (Chapters 4–15).

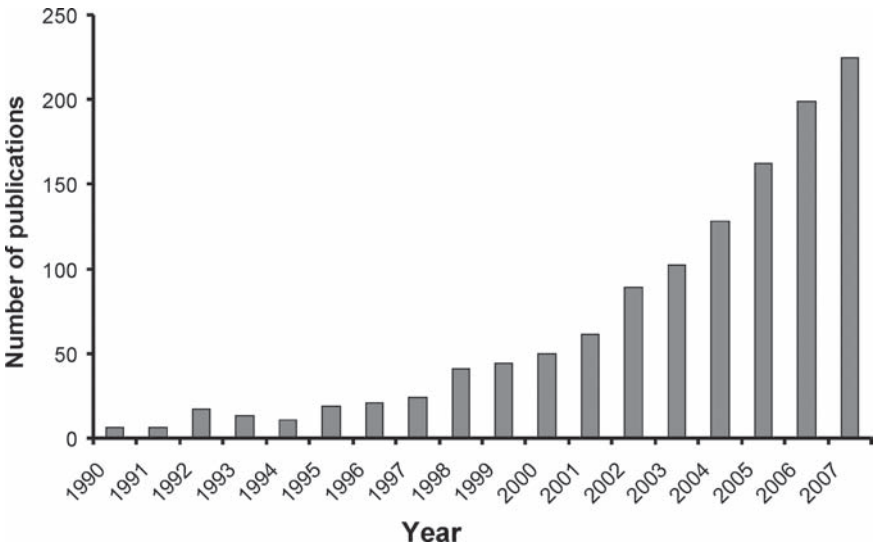

Fig. 1.1 Number of publications related to functional nucleic acids sensors. Results obtained using SciFinder on December 1, 2007. Search terms: sensor and (ribozyme, or deoxyribozyme, or DNAzyme, or aptamer or aptazyme)

Chapter 4, by Tan and coworkers, provides snapshots of recent progress in fluorescent aptamer sensors. Fluorescent sensors have the advantage of being highly sensitive. Many fluorophores can be detected at concentrations less than 1 nM with common benchtop fluorimeters. Many portable fluorimeters are also commercially available, which allows the use of fluorescent sensors in the field. Fluorescence methods give immediate response and can be used in real time Several fluorescence-based signal transduction schemes, such as direct fluorophore labeling, fluorescence resonance energy transfer (FRET), fluorescence quenching, fluorescence anisotropy, and light-switching excimers, are discussed. Examples given have shown that the aptamer sensors can result in simple, rapid, sensitive, and selective biological assays of a wide range of molecules, including organic and protein molecules and tumor cells.

Our ability to create allosteric nucleic acid enzymes (aptazymes) by rational design or in vitro selection has considerably expanded the utility of ribozymes and DNAzymes. Li and Chiuman discuss progress in fluorescent ribozyme and deoxyribozyme sensors in Chapter 5. They begin their chapter by reviewing several strategies that have been used to produce ribozymes or DNAzymes which can generate fluorescent signals upon catalysis. This part is followed by discussion of the efforts in their own laboratory in creating fluorescence-signaling and RNA-cleaving deoxyribozymes by in vitro selection. The third part of their chapter reviews several published approaches that can link the binding action of an aptamer to catalytic activity of a fluorescence-signaling ribozyme or deoxyribozyme.

Recent development in nanotechnology has brought new opportunity in FNA-sensing applications, as nanomaterials possess strong and unique signaling properties for FNA binding of targets and catalysis. FNA-modified nanomaterials are ideally suited for analytical applications. Chapter 6, by Liu and Lu, first discusses simple colorimetric sensors that combine DNAzymes, aptamers, and aptazymes with metallic nanoparticles. Reporting systems obtained from modifying quantum dots with FNAs for fluorescent sensing are also presented. Their chapter closes with an example in which colorimetric sensors are formulated into a “dipstick” test, making it possible for simple on-site real-time detection and quantification without instrumentation.

Electronic devices can be routinely miniaturized in terms of size and power consumption and can usually offer ultrahigh sensitivity, good reproducibility, and fast response; this may explain why electrochemical sensing exploiting FNAs has become a very active research area in recent years. In Chapter 7, Xiao and Plaxco review some exciting sensors that exploit the unusual physical properties of aptamers, properties which render them uniquely well suited for application to impedance- and folding-based electrochemical sensors. These sensors make use of electrode-bound aptamers that undergo reversible, binding-induced folding, which translates into an electronic signal. These sensors can be used to provide reagentless and highly sensitive detection of specific biomolecules in blood serum and other complex, interference-ridden sample matrices.

One of the most important parameters in analytical applications is the detection sensitivity. Although many of the FNAs possess high affinity for their targets, methods

that can further enhance sensitivity of FNA sensors are critical to their wide applicability. In Chapter 8, Willner and coworkers review many interesting methods for the design of FNA sensors with signal amplification capabilities.

In addition to sensing applications, FNAs have been shown to be quite useful in many other analytical applications. This book covers several promising applications such as affinity separation, high-throughput screening, and DNA-based computing. New sensing formats such as aptamer arrays, solid-phase assays, and FNA-based nanostructures provide excellent opportunity for aptamer researchers to take full advantage of FNA-based sensors for a broader range of applications.

In Chapters 9 and 10, the applications of aptamers for affinity separations have been reviewed using capillary electrophoresis (CE) and stationary separations, respectively. In the area of CE-based affinity separation, Guthrie and Le categorize the methods into competitive and noncompetitive assays, fluorescence polarization assays, nonequilibrium capillary electrophoresis of equilibrium mixtures, and affinity-polymerase chain reaction-CE assays. CE not only provides a sensitive detection method but can also be used as a convenient system to determine important parameters of aptamer–target complexation (such dissociation rate and equilibrium binding constants and binding stoichiometries). In Chapter 10, Ravelet and Peyrin provide a discussion on recent activities in using aptamers as target-specific ligands for the separation and capture of various analytes in affinity chromatography and related affinity-based methods such as magnetic bead technology.

The increasingly wider applications of aptamers demand the creation of more aptamers for new targets and the use of many different aptamers for simultaneous detection and quantifications of multiple biological targets. In Chapter 11, Ellington and colleagues discuss recent development both in high-throughput, automated aptamer selection and in construction of aptamer microarrays. For example, the results from their study have shown that arrayed aptamers are almost as sensitive as their solution-phase counterparts and, when assembled together, can provide both specific and general diagnostic signals for proteins and other analytes. Coupling of immobilized aptamers to amplification methods such as those reviewed in Chapter 8 can greatly enhance detection signals.

FNA-based microarrays and other solid-phase applications underscore the importance of developing methods that can immobilize FNAs onto solid supports and maintain the activity of layered FNAs. In Chapter 12, Brennan and coworkers first provide a good overview of methods that can be explored for immobilization of FNAs. This review is followed by discussion of their published efforts in which they studied sol-gel encapsulation of FNAs for the development of solid-phase fluorimetric assays for biosensing and proteomics.

Aptamers and nucleic acid enzymes not only have the potential to be directly exploited as therapeutic agents but can also be explored as tools to derive small-molecule drugs. In Chapter 13, Rentmeister and Famulok have provided a review on research activities where FNA-based sensors have been explored as reporting tools to facilitate high-throughput screens (HTS) to search for small-molecule probes or inhibitors for protein targets that may not have a convenient assay to conduct HTS.

In Chapter 14, Macdonald and Stojanovic discuss an interesting line of research that explores FNAs for DNA-based computation. For example, they have produced various molecular logic gates using RNA-cleaving DNAzymes as switches and single-stranded oligonucleotides as inputs and outputs. These logic gates can be further combined to produce basic computational circuits such as half- and full adders. Their DNAzyme-based logic gates can also be assembled into automata to perform complex computational tasks such as game playing. Their efforts may eventually lead to applications in which FNAs can be utilized as autonomous diagnostic and/or therapeutic molecular devices.

Finally, in Chapter 15, Mao and coworkers provide a review on some of their work on exploring DNA-based enzymes to create interesting DNA assemblies that can function as molecular machines. These DNA machines are regarded as "nanomachines" because their sizes are smaller than 100 nanometer (nm). For example, in Mao's lab, some specific RNA-cleaving DNAzymes are combined with matching substrates and other regulatory DNA strands in unique ways to create autonomous nanomotors, which can extract chemical energy from RNA substrates and convert it into a mechanical motion.

Work reviewed in all the chapters of this book, although quite exciting, represents only a portion of demonstrated examples where FNAs are applied as analytical tools. We hope that by presenting the state-of-the-art methods and technologies concerning FNAs, readers can gain appreciation of what have been accomplished and what remains to achieve. Given the rapid progress made recently in FNAs, we are confident that many of the FNA sensors and emerging analytical applications will be commercially available in the near future. However, many challenges still remain to be tackled before we can realize the full potential of FNAs. Therefore, we hope the information presented in this book can function as an "inducer" to engage more researchers to link their research activities to the expansion of the aptamer arena. We also hope that the demonstrated examples and remaining challenges discussed in this book can motivate future generations of scientists to select this field as their interest of study.

Following the groundbreaking discovery of ribozymes, it took merely 10 years for a group of scientists to come up with the brilliant idea of conducting selection and evolution experiments using DNA or RNA in test tubes. This great invention has now become a powerful vehicle for molecular scientists to travel through the land of nucleic acids for more discoveries and applications. We hope this book can serve as a useful guide for those who would like to take such a journey to search for new frontiers in the amazing "jungle" of functional nucleic acids.

We thank all the authors for their efforts in putting together the excellent chapters and Jack Szostak for writing the preface of the book. We also want to thank the members of our two laboratories for either their assistance or their patience with us while we were working on this project. Special credits go to Juewen Liu and William Chiuman for their assistance in chapter writing and other important issues such as copyright permission acquisition, and Janet Sinn-Hanlon of Imaging Technology Group at the Beckman Institue and Jung Heon Lee and Zidong Wang of Department of Materials Sciences and Engineering at the University of Illinois at Urbana-Champaign, in design of the cover graphics, and

proofreadings. We want to express our gratitude to Kenneth Howell (Springer) and Radislav Potyrailo (GE Global Research Center) for the invitation to assemble this book and for their guidance along the way. Last, but not least, we thank our families for their support and patience.

References

1. Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E. and Cech, T.R. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. Cell 31:147–157.
2. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell 35:849–857.
3. Ellington, A.D. and Szostak, J.W. (1990) In vitro selection of RNA molecules that bind specific ligands. Nature (Lond.) 346:818–822.
4. Robertson, D.L. and Joyce, G.F. (1990) Selection in vitro of an RNA enzyme that specifically cleaves single-stranded DNA. Nature (Lond.) 344:467–468.
5. Tuerk, C. and Gold, L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science 249:505–510.
6. Lee, J.F., Hesselberth, J.R., Meyers, L.A. and Ellington, A.D. (2004) Aptamer database. Nucleic Acids Res. 32:D95–D100.
7. Thodima, V., Piroozina, M. and Deng, Y. (2006) RiboaptDB: a comprehensive database of ribozymes and aptamers. BMC Bioinformatics 7(suppl 2):S6.
8. Breaker, R.R. (2004) Natural and engineered nucleic acids as tools to explore biology. Nature (Lond.) 432:838–845.
9. Fedor, M.J. and Williamson, J.R. (2005) The catalytic diversity of RNAs. Nat. Rev. Mol. Cell. Biol. 6:399–412.
10. Achenbach, J.C., Chiuman, W., Cruz, R.P. and Li, Y. (2004) DNAzymes: from creation in vitro to application in vivo. Curr. Pharm. Biotechnol. 5:321–336.
11. Joyce, G.F. (2004) Directed evolution of nucleic acid enzymes. Annu. Rev. Biochem. 73:791–836.
12. Silverman, S.K. (2005) In vitro selection, characterization, and application of deoxyribozymes that cleave RNA. Nucleic Acids Res. 33:6151–6163.
13. Mandal, M. and Breaker, R.R. (2004) Gene regulation by riboswitches. Nat. Rev. Mol. Cell Biol. 5:451–463.
14. Tucker, B.J. and Breaker, R.R. (2005) Riboswitches as versatile gene control elements. Curr. Opin. Struct. Biol. 15:342–348.
15. Winkler, W.C. (2005) Riboswitches and the role of noncoding RNAs in bacterial metabolic control. Curr. Opin. Chem. Biol. 9:594–602.
16. Winkler, W.C. and Breaker, R.R. (2005) Regulation of bacterial gene expression by riboswitches. Annu. Rev. Microbiol. 59:487–517.
17. Sullenger, B.A. and Gilboa, E. (2002) Emerging clinical applications of RNA. Nature (Lond.) 418:252–258.
18. Nimjee, S.M., Rusconi, C.P. and Sullenger, B.A. (2005) Aptamers: an emerging class of therapeutics. Annu. Rev. Med. 56:555–583, 553 plates.
19. Lee, J.F., Stovall, G.M. and Ellington, A.D. (2006) Aptamer therapeutics advance. Curr. Opin. Chem. Biol. 10:282.
20. Famulok, M., Hartig, J.S. and Mayer, G. (2007) Functional aptamers and aptazymes in biotechnology, diagnostics, and therapy. Chem. Rev. (Washington, DC) 107:3715–3743.
21. Cho, E.J., Rajendran, M. and Ellington, A.D. (2005) Aptamers as emerging probes for macromolecular imaging. Top. Fluoresc. Spectrosc. 10:127–155.

22. Famulok, M. (2005) Allosteric aptamers and aptazymes as probes for screening approaches. Curr. Opin. Mol. Ther. 7:137.
23. Blank, M. and Blind, M. (2005) Aptamers as tools for target validation. Curr. Opin. Chem. Biol. 9:336–342.
24. Famulok, M. and Mayer, G. (2005) Intramers and aptamers: applications in protein-function analyses and potential for drug screening. ChemBioChem 6:19–26.
25. Ravelet, C. and Peyrin, E. (2006) Recent developments in the HPLC enantiomeric separation using chiral selectors identified by a combinatorial strategy. J. Separ. Sci. 29:1322–1331.
26. Lu, Y. and Liu, J. (2007) Smart nanomaterials inspired by biology: dynamic assembly of error-free nanomaterials in response to multiple chemical and biological stimuli. Acc. Chem. Res. 40:315–323.
27. Lu, Y. and Liu, J. (2006) Functional DNA nanotechnology: emerging applications of DNAzymes and aptamers. Curr. Opin. Biotechnol. 17:580–588.
28. Katz, E. and Willner, I. (2004) Nanobiotechnology: integrated nanoparticle–biomolecule hybrid systems: synthesis, properties, and applications. Angew. Chem. Int. Ed. 43:6042–6108.
29. Feldkamp, U. and Niemeyer, C.M. (2006) Rational design of DNA nanoarchitectures. Angew. Chem. Int. Ed. 45:1856–1876.
30. Famulok, M., Mayer, G. and Blind, M. (2000) Nucleic acid aptamers: from selection in vitro to applications in vivo. Acc. Chem. Res. 33:591–599.
31. Rajendran, M. and Ellington, A.D. (2002) Selecting nucleic acids for biosensor applications. Comb. Chem. High Throughput Screen. 5:263–270.
32. Hesselberth, J., Robertson, M.P., Jhaveri, S. and Ellington, A.D. (2000) In vitro selection of nucleic acids for diagnostic applications. Rev. Mol. Biotechnol. 74:15–25.
33. Breaker, R.R. (2002) Engineered allosteric ribozymes as biosensor components. Curr. Opin. Biotechnol. 13:31–39.
34. Lu, Y. (2002) New transition metal-dependent DNAzymes as efficient endonucleases and as selective metal biosensors. Chem. Eur. J. 8:4588–4596.
35. Navani, N.K. and Li, Y. (2006) Nucleic acid aptamers and enzymes as sensors. Curr. Opin. Chem. Biol. 10:272–281.
36. Yang, L. and Ellington, A.D. (2006) Prospects for the de novo design of nucleic acid biosensors. Fluoresc. Sens. Biosens. 5:41, 43.

Chapter 2
Natural Functional Nucleic Acids: Ribozymes and Riboswitches

Renaud Tremblay, Jérôme Mulhbacher, Simon Blouin, J. Carlos Penedo, and Daniel A. Lafontaine

Abstract Natural functional nucleic acids are of primary importance in most cellular processes. Although artificial RNA motifs with functional properties can routinely be generated in research laboratories, the efficiency of their naturally occurring counterparts is hardly matched. Natural ribozymes and riboswitches are examples of Nature's prowess at creating exceedingly good catalysts and ligand-sensing aptamers. This review focuses on natural ribozymes and riboswitches and attempts to highlight how RNA can rival proteins when it comes to show off its capabilities.

2.1 Introduction

RNA molecules are involved in countless essential reactions in living cells – in addition to their roles in information storage and architectural framework, they also perform catalytic reactions, as best exemplified by the peptidyl transfer reaction of ribosomal RNA.[1] However, this nowadays widely accepted notion has not always been the case. Indeed, only 25 years ago, RNA was only considered to be the messenger that would carry genetic information from DNA to the ribosome for the production of proteins. Protein enzymes were considered to be the only true "molecular effectors," which participated in many important biological processes. However, when Thomas Cech[2] and Sidney Altman[3] simultaneously discovered that RNA can exhibit catalytic properties, similar to protein enzymes, it changed the way scientists envisioned how RNA is involved in molecular reactions. These RNA enzymes, the so-called ribozymes, playing both informational and catalytic roles, inspired the "RNA world" hypothesis. This proposal was put forward to address the "chicken-and-egg" problem of how a translation system could evolve without proteins already in place. Thus, if RNA molecules, or some other

R. Tremblay, J. Mulhbacher, S. Blouin, and D.A. Lafontaine
Université de Sherbrooke, Département de Biologie
daniel.lafontaine@usherbrooke.ca

J.C. Penedo
University of St. Andrews, School of Physics and Astronomy
jcp10@st-andrews.ac.uk

Y. Li and Y. Lu (eds.), *Functional Nucleic Acids for Analytical Applications*,

RNA-related biopolymers, could encode information and perform catalysis at the same time, then protein-based processes could well have evolved from those giving rise to the complexity and variety of modern cellular processes. If this is true, it seems plausible to suggest that molecular remnants of that RNA era would still remain today.

A few years ago, RNA came into the spotlight for a second time. Indeed, it was found that RNA was able to use feedback regulatory loops, such as retro-inhibition, to monitor and control its own expression – all of this apparently in the absence of any helper protein. These molecules, the so-called riboswitches, are literally "molecular switches" that can sense the presence of small cellular metabolites and, upon binding to them, switch into a different conformation to either promote or inhibit transcription or translation of the encoded protein. The first examples of this regulatory behavior were found studying the regulation of biosynthetic enzymes producing vitamins B_1, B_2, and B_{12}.[4–8] Because riboswitches can operate in the absence of proteins in vitro, it has been proposed that they could be remnants of the RNA world that was essentially RNA based.[9] Whether riboswitches really work in the absence of protein in the cell still remains elusive, and this point will most probably require additional investigations.

For a primitive world to be viable, it inherently requires that a large array of catalytic reactions and metabolic regulations are performed by RNA molecules. Thus, to probe to what extent RNA is a "good contortionist" at performing various catalytic reactions, a procedure has been put forward that can be described as an accelerated evolution which is done in a test tube. This procedure, called SELEX (systematic evolution of ligands by exponential enrichment),[10–12] relies on the starting hypothesis that, in a given population, at least one RNA sequence exists that is able to perform the desired function. Following this approach, artificially engineered ribozymes have been isolated for a variety of reactions such as hydrolysis, RNA transesterification, peptidyl transfer, ester hydrolysis, isomerization, and Diels–Alder cyclo-addition (recently reviewed[13]). Thus, although RNA has far fewer functional groups compared to proteins, it can still perform a large set of chemical reactions. SELEX has also been used to test RNA molecules for their ligand-binding capacity (recently reviewed[14]). By employing ligands bound to a solid phase, many artificial aptamers have been isolated that respond to various metabolites such as theophylline and adenosine triphosphate (ATP). A natural extension to this was the creation of artificial ribozymes that could be controlled by external signals. Indeed, many allosterically controlled ribozymes were created that could respond to small organic compounds, proteins, pH, light, etc.[14–16] Although a large number of artificial ribozymes and aptamers have been isolated, this review focuses on the natural representatives and attempts to emphasize the variety of tricks that RNA has up its sleeve to rival protein activities. The review of artificially derived functional nucleic acids is provided in Chapter 3 of this volume.

2.2 Ribozymes

RNA catalysis was very likely crucial for the development of early life on Earth. Ribozymes are thought to be "molecular fossils" still in place today. Natural ribozymes perform various enzymatic reactions, and they are usually classified by

their relative size. The large ribozymes [>300 nucleotides (nt)] are constituted of the self-splicing introns (groups I and II) and the protein-assisted RNase P (Fig. 2.1). Although self-splicing introns perform transesterification reactions, the RNase P employs a hydrolysis mechanism to process the 5′-end of transfer RNAs, thus making the reaction practically irreversible. Several crystal structures are available for the group I intron that reveal different steps of the chemical reaction.[17–20] In all cases, the core domain is well conserved and in agreement with previous biochemical data (recently reviewed[21]). Strikingly, a guanine suggested almost 20 years ago[22] to be involved in a triple-base interaction is observed in all crystals. This interaction is also consistent with data from Bass and Cech reporting that the *Tetrahymena* intron is able to use guanosine analogues.[23]

In contrast to other natural ribozymes, RNase P is a multiple turnover enzyme that recognizes and cleaves its substrate in *trans* through recognition of secondary and tertiary structural elements in the substrate.[24] Moreover, RNase P is able to process various structurally different substrates such as pre-tRNAs, 4.5 S RNA, bacteriophage 80-induced RNA, the mRNA from the polycistronic his operon, transient structures adopted by riboswitches, and other cellular substrates.[25] The RNA component of RNase P is strongly conserved, which suggests its importance for cellular metabolism.[26] Evidence of this importance has been recently found on an eukaryotic RNase P, whose RNA moiety is active in absence of proteins, akin to its prokaryotic counterpart.[27] Complete RNase P RNAs from bacterial sources as well as isolated domains have recently been crystallized.[26,28–30] The two complete RNase P crystals were solved for types A and B, which have structurally distinct secondary structures.[25] Unfortunately, neither crystal is able to deliver insights into the conformation of the active structure because both have some associated technical problems.[25] Nevertheless, both structures exhibit a compact fold with a flat surface

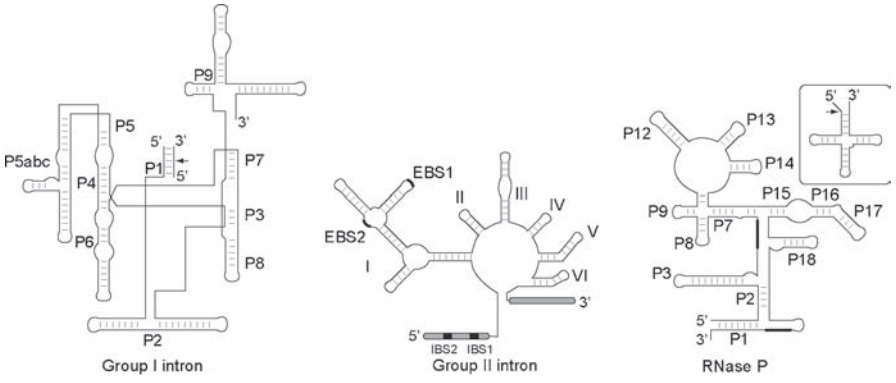

Fig. 2.1 Naturally occurring large ribozymes. Secondary structures are shown for each species. Cleavage sites are shown by *arrows* for the group I intron and the RNase P. The intron and exon binding sequences, *IBS* and *EBS*, respectively, are shown for the group II intron ribozyme. *Gray regions* indicate the exons of the group II intron ribozyme. A tertiary interaction is shown by *bold bars* for the RNase P. The substrate of the RNAse P, a tRNA, is shown in the *inset*

to which a pre-tRNA binds. The nonconserved regions are located away from the binding site and participate in long-range tertiary interactions most probably important for the global folding of the ribozyme. However, even with these crystal structures in hand and together with the huge amount of biochemical data, we are still far from understanding the RNAse P structure and the details of RNA–protein interactions that are part of it. Nevertheless, despite their inherent lack of precision, tertiary models based on isolated crystallized domains of the RNase P and biochemical results provide stimulation for new testable hypotheses.[25]

A mechanistically distinct class of smaller ribozymes also exists that includes the hammerhead, hairpin, hepatitis delta virus (HDV), Varkud satellite (VS), and glmS ribozymes (Fig. 2.2). All these ribozymes perform a self-cleaving transesterification, generating a hydroxyl and a cyclic phosphate terminus that can then be used to catalyze the reverse reaction according to a ligation mechanism. Crystal structures have been obtained for all small ribozymes,[31–37] excluding the VS ribozyme, for which a tertiary model has been built using a combination of comparative gel electrophoresis and fluorescence resonance energy transfer (FRET) distance restraints.[38,39] Because they can function in the absence of divalent ions, in contrast to large ribozymes, it has been suggested that small ribozymes might employ a nucleobase-driven catalysis.[40,41] A new ribozyme seems to have been identified within the human β-globin mRNA that is involved in the cotranscriptional processing.[42] Interestingly, the cleavage site is contained within a region predicted to exhibit a structure similar to the one observed for the hammerhead

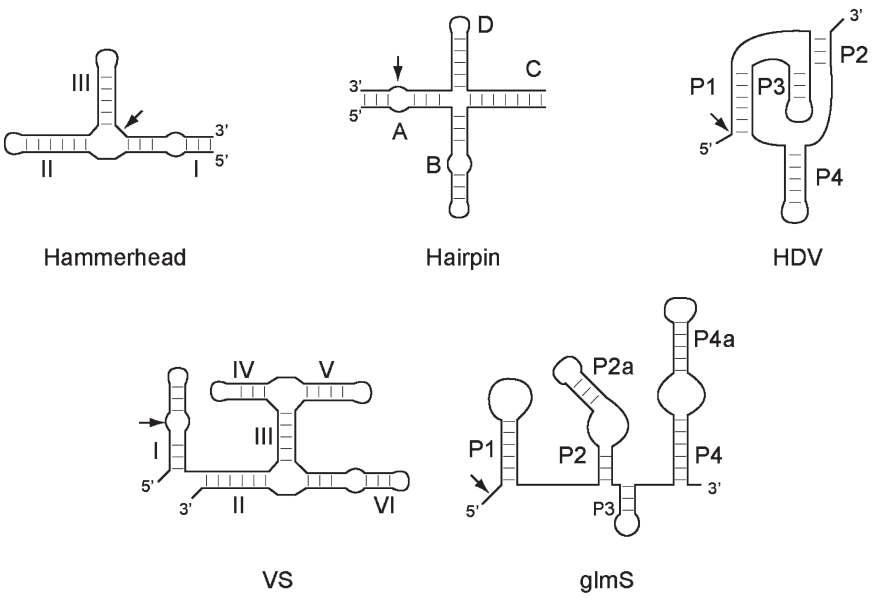

Fig. 2.2 Small nucleolytic ribozymes. Secondary structures and cleavage sites are shown for each ribozyme. The *Schistosoma* variant is represented for the *hammerhead* ribozyme

ribozyme. However, this similarity is only superficial given that the product termini are more related to those generated by the RNase P. Moreover, in addition to providing a model system to study catalysis, small ribozymes are also intensively employed in gene therapy to reduce the expression of specific mRNAs.[43] Although this avenue is promising, various technical problems such as the RNA stability and gene targeting must be addressed to fully exploit this concept.

Natural ribozymes are heavily biased toward transesterification reactions. However, crystallization of the big subunit of the ribosome made clear that the peptidyl transfer reaction that catalyzes the ribosome is very likely performed exclusively by the RNA component, making the ribosome, in fact, a ribozyme.[44] In this case, it has been proposed that a particular adenosine, A2451 in *Escherichia coli*, is involved in a proton transfer reaction during the catalytic step.[1] Interestingly, a peptidyl transferase ribozyme was previously isolated through SELEX,[45] giving support to the notion that the RNA moiety of the ribosome is directly responsible for the catalytic activity, and that ribosomal proteins could function more or less as folding chaperones. Using single-molecule fluorescence spectroscopy, structural information has recently been obtained concerning tRNA dynamics on the ribosome during translation.[46,47] In these studies, it was demonstrated that dynamics of ribosome function are very important for the translation process, and that the ribosomal recognition of correct codon-anticodon pairs drives ribosomal rotational movement during tRNA selection, which is most probably important for the proofreading process.

Various excellent reviews on ribozymes have been published recently.[21,48–54] Here, we focus on the hammerhead ribozyme because it is the one for which a larger body of work on the catalytic mechanism, structure, and folding has been accumulated over the years. Also, a recent crystal structure[36] has shed light on years of discrepancy between structure and function data. Particular attention is also devoted to the VS ribozyme, which is the only natural small ribozyme for which a crystal structure is yet to be solved.

2.2.1 The Hammerhead Ribozyme

The hammerhead is a catalytic motif found in RNA genomes of numerous plant pathogens, and it is involved in the processing of RNA issued from the rolling-circle amplification.[55] For more than 20 years, the hammerhead has been defined as a single three-way junction organized around a conserved core containing the catalytic center (see Fig. 2.2). This autocatalytic species performs the reversible site-specific cleavage of its backbone via a transesterification reaction in which the 2′-oxygen attacks the adjacent 3′-phosphorus, ultimately leading to 5′-hydroxyl and 2′,3′-cyclic phosphate termini. The reaction proceeds with inversion of the configuration at the phosphorus center, indicating that the reaction occurs by transesterification via an S_N2 mechanism. The catalytic reaction is accelerated by at least 10^6-fold compared to the nonenzymatic reaction.[56] The hammerhead

can be engineered as a true enzyme by performing a nick in one of the capping loops, effectively generating a substrate and an enzyme strand. The bimolecular hammerhead catalytic reaction can thus be viewed as follows: association between the substrate and the ribozyme, chemical cleavage step, and dissociation of the ribozyme–products complex. It has been deduced that the substrate association and product dissociation steps have kinetic and thermodynamic properties very similar to an RNA helix-coil transition, and that the chemical step of the cleavage reaction is almost identical throughout all hammerhead ribozymes.[57] In most conditions, the cleavage is at least 100-fold faster than the ligation rate.[58,59]

The hammerhead ribozyme is one of the most studied small catalytic RNAs, and recent advances have shown that, for years, studies were performed on a minimal active form lacking an important loop–loop interaction.[60,61] Indeed, because, of the "RNA reductionism" approach, where an RNA of interest is truncated to its smallest active size, RNA researchers have inadvertently removed a very important loop–loop interaction between stems I and II (Fig. 2.3). It appears that the presence of this tertiary interaction decreases the concentration of divalent ions required for catalytic activity by almost 100-fold compared to the minimal form. This unexpected finding ended a long-standing discrepancy between physiological divalent ions concentrations and that required for significant in vitro activity of the minimal form. Thus, it is important to realize that experiments conducted on the minimal form (i.e., without the loop–loop interaction) might not reflect what

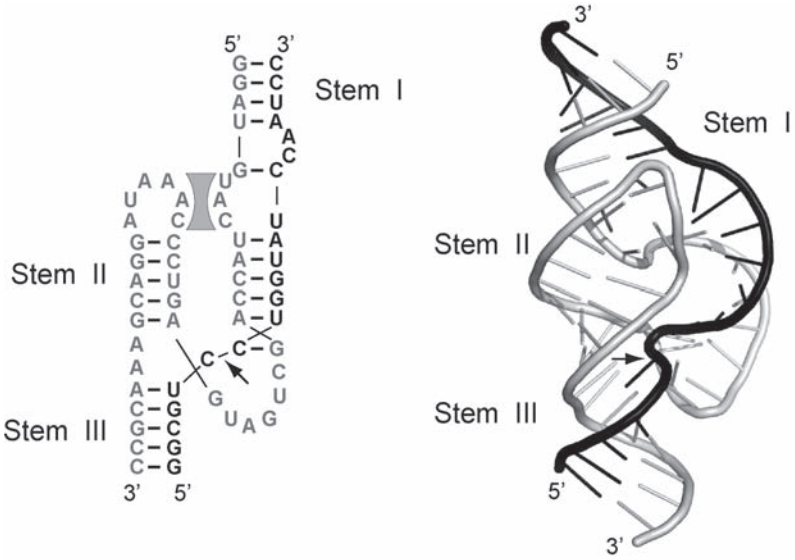

Fig. 2.3 The *Schistosoma* hammerhead ribozyme. The secondary (*left*) and the tertiary (*right*) structures are shown. The substrate and the ribozyme moieties are colored in *gray* and in *black*, respectively. A tertiary interaction and the cleavage site are indicated by the *shaded region* and an *arrow*, respectively[36]

is occurring in the natural sequence. Nevertheless, we believe that a comparison between the biochemical data of the minimal and natural forms of the hammerhead is worthwhile to highlight the importance of what initially were taken as noncrucial structural elements and also to emphasize the need for caution when taking a "reductionist approach."

In its minimal form, the folding and dynamics of the hammerhead have been studied using transient electric birefringence,[62] comparative gel electrophoresis,[63] FRET,[64–66] 2-aminopurine fluorescence,[67,68] binding of bases to abasic sites,[69] cross-linking,[70–73] and solution NMR experiments.[74–76] In agreement with these results, an ion-induced folding scheme of the ribozyme has been elaborated where the folding occurs by a sequential two-stage process, in which both are induced by the noncooperative binding of divalent metal ions.[65,77] During the first step, a series of purine base pairs are formed to produce the domain 2 region, and this occurs in the micromolar (μM) range of magnesium ion concentration, resulting in the coaxial stacking of helices II and III. The second folding transition occurs in the millimolar (mM) range and involves the formation of an uridine turn (domain 1), which brings helix I close to helix II. Because the activity of the ribozyme is observed in the millimolar range of magnesium, this folding step is associated with the active form, as seen in crystal structures of the minimal ribozyme form.[31,32]

The loop–loop interaction was recently discovered, and it was shown that its presence was associated with hammerhead ribozymes active at very low magnesium ion concentrations.[78–80] It is likely that the tertiary interaction restricts internal movements of loops I and II, thus stabilizing states closely resembling the active conformation, consistent with the reduced required magnesium ions. Indeed, the effect of the presence of the loop–loop interaction in the folding pathway of the ribozyme has been studied using FRET analysis. In this study, it was found that the ion-induced folding is greatly enhanced in the presence of the loops promoting the formation of the active conformation in a single step at micromolar magnesium ion concentrations[81]; this is consistent with the loop–loop interaction helping in the close juxtaposition of loops I and II, which lowers the required magnesium ion concentration to achieve the native state. Consistent with this view, recent studies have observed that, under low ionic strength conditions, divalent metal ions appear to stimulate folding of this ribozyme species.[82,83] The loop–loop interaction can thus be seen as a "folding enhancer" element[81] that is essential for the achievement of catalytic activity under a physiologically low magnesium concentration.

The newly solved *Schistosoma mansoni* hammerhead[84] crystal structure[36] clearly illustrates how the loop–loop interaction is involved in the formation of the global structure of the ribozyme (see Fig. 2.3). In their work, Martick and Scott found that the global folding of the molecule primes the ribozyme for catalysis, and that the structure of the core of the *Schistosoma* hammerhead differs from the structures reported for minimal configurations.[36] Because of multiple structural rearrangements, the core of the ribozyme can no longer be considered as having two discrete domains.[85] Instead, a single complex network of interactions span both domains I and II, and in addition, the scissile bond is readily positioned for an in-line attack from the 2′-OH, which is what we expect from a catalytic RNA molecule.

These observations indicate that in addition to altering the global folding of the hammerhead, the presence of the loop–loop interaction modifies the local arrangement of the catalytic core. This concept was emphasized in a recent review where it was attempted to explain the cleavage rates of about 50 minimal hammerhead from available crystallographic data.[56] Half the results could not be readily explained in terms of the structure, highlighting a strong discrepancy between available biochemical and structural data.[85] On the other hand, the *S. mansoni* hammerhead structure is in excellent agreement with available biochemical data and should thus be a reasonable approximation of an active molecule. In the crystal, G12 and the 2′-OH of G8 are positioned very close to the scissile bond, suggesting their roles in an acid–base catalysis. In support of this, a recent study performed on a minimal hammerhead molecule showed the importance of G8 and G12 for the pH profile of the catalytic reaction.[86] In general, the new crystal structure explains most of the previously irreconcilable sets of published results, providing a consistent model.

Interestingly, a similar situation has previously occurred in the context of another small catalytic RNA, the hairpin ribozyme (see Fig. 2.2). The hairpin ribozyme is involved in the cleavage and ligation of the negative strand of sTRSV (tobacco ringspot virus satellite). The minimal sequence of the hairpin ribozyme consists of two stems connected via a hinge region.[87] Both stems include an internal region that contains loops A and B, with the former carrying the substrate cleavage site. However, in the natural sequence, the ribozyme folds as a four-way junction[88] where the presence of the junction is associated with a requirement approximately 1,000-fold lower for magnesium ions to achieve efficient activity.[89] In addition, time-resolved and single-molecule FRET studies have demonstrated that the junction is responsible for the stabilization of the active ribozyme form[90] and the acceleration of the folding via a discrete intermediate.[91] Thus, the loop–loop interaction and the four-way junction in the hammerhead and hairpin ribozymes, respectively, are folding elements that help the RNA molecule to achieve the final folded state under physiological conditions.

2.2.2 The Varkud Satellite Ribozyme

The VS ribozyme is found in the 881 nt VS RNA located in the mitochondria of *Neurospora*, which is transcribed from the Varkud satellite DNA.[92] Although the basic chemistry of the cleavage reaction appears to be close to that of other small ribozymes, the VS ribozyme differs significantly in a number of respects. In addition to being the largest ribozyme and having a secondary structure quite different from other ribozymes, the formation of the ribozyme–substrate complex is not performed via an extensive stretch of Watson–Crick base pairs (see Fig. 2.2). Instead, a loop–loop interaction mediates the formation of the complex, indicating that substrate recognition is mainly achieved via tertiary interactions.[93] This loop–loop interaction has been shown to induce a change in the secondary structure of the substrate by which a number of base pairs are rearranged to yield a “shifted”

substrate conformation,[94] which is most probably important for the creation of an environment favorable to catalysis. Even if the natural configuration of this catalytic RNA is in *cis*, where the substrate is covalently linked to the ribozyme sequence, *trans*-cleavage activity is very efficient and the system can thus be viewed as a bimolecular one, where the substrate is a single stem-loop (stem I), and the ribozyme is composed of five helices (stems II–VI). The ribozyme is organized around two three-way junctions that are arranged as an H shape (see Fig. 2.2). Although a crystal structure of this ribozyme has yet to be obtained, a great amount of biochemical and structural data has been accumulated over the years, from which a detailed catalytic model has been proposed.[95,96]

The substrate contains an asymmetrical internal loop in which the scissile bond is located (see Fig. 2.2). Several NMR studies have found that in the ground state of substrate (i.e., not complexed to the ribozyme), the internal loop is composed of two sheared G•A base pairs and a protonated A^+•C base pair.[97,98] However, by using mutations to obtain a "shifted" substrate, it has been shown that one of the sheared base pairs is disturbed, creating alternative interactions in the internal loops, and additional magnesium ion binding sites.[99] It is clear that additional structural studies on a ribozyme–substrate complex are much needed to fully understand how the ribozyme participates in both the substrate reorganization and the catalytic steps.

Because of the large size of the ribozyme and the lack of available crystal structure, it has been difficult to obtain three-dimensional information describing the global fold of the molecule. Nevertheless, indirect experimental evidence was obtained using biochemical approaches that were used to build a general model. One of the first observations was the discovery of a loop–loop interaction between the substrate and the ribozyme,[93] which effectively places loops I and V in close juxtaposition. A second constraint was revealed by a short-wavelength UV cross-link between two adjacent stems (II and VI) suggesting their close proximity. These results, together with a body of biochemical work, inspired a model for the arrangement of the ribozyme and suggested the importance of the A730 loop.[100–103]

Using a combination of comparative gel electrophoresis and FRET, the global structure of each junction was determined.[38,39] Both junctions were found to undergo coaxial stacking of two helices by the noncooperative binding of magnesium ions. In the 2–3–6 junction, helices III and VI are coaxially stacked, with an acute angle subtended between helices II and VI.[38] Surprisingly, a three-way junction of very similar sequence was found in the *Haloarcula marismortui* 23 S rRNA.[44] When transplanted into the VS ribozyme, an efficient cleavage activity is retained together with complete cleavage of the substrate, indicating that both three-way junctions most probably adopt a very similar global structure. The core of the VS junction was thus modeled using the ribosomal junction. However, no junction similar to the 3–4–5 VS junction was found in the ribosome, leaving uncertainty about the local structure of the VS junction. In addition, because the two VS junctions share a common helix (stem III), a novel comparative gel assay was designed to determine their relative orientation. By extending stems II and V, the amplitude of the end-to-end distance between each stem was magnified and their dihedral angle was thus determined to be about 75°.[39] Although the resulting model does not include base

bulges and internal loops (Fig. 2.4), it still provides an excellent starting point for structure–function analysis of the VS ribozyme.

Using the model described above and a few natural constraints, the location of the substrate can be deduced. First, the substrate is naturally attached to the ribozyme by the 5′-extremity of stem II, and based on previous biochemical analysis, the substrate loop is involved in a tertiary interaction with the loop of stem V.[93,100] Based on these two spatial constraints, it appears logical to position the substrate between stems II and V (see Fig. 2.4). By doing so, the internal loop of the substrate, and hence the scissile bond, is brought in close proximity to the A730 loop located in stem VI. Based on this, it is possible to envision that the A730 loop is directly involved in the catalysis. This hypothesis is in agreement with an extensive mutational analysis which has shown that most of the nucleotide sequence of the VS ribozyme plays an important architectural role, but in contrast, that the A730 loop is most probably an important component of the active site.[104] Moreover, hydroxyl radical probing[105] and nucleotide analogue interference mapping (NAIM)[106] data have revealed that stem II is important for substrate binding, which is in agreement with the proposed model. The substrate is thus thought to

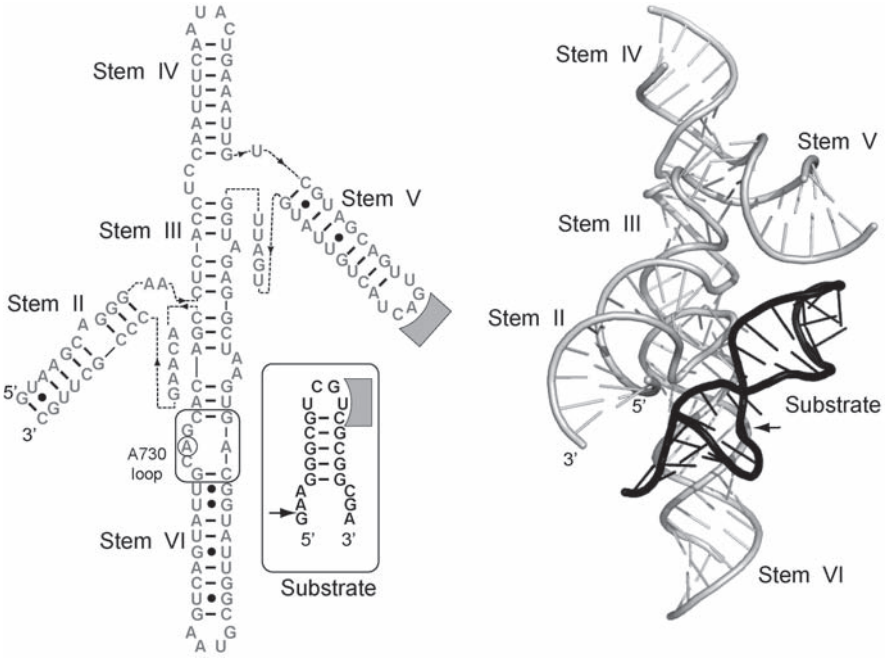

Fig. 2.4 The Varkud satellite (VS) ribozyme. The secondary structure (*left*) is shown to better represent the tertiary arrangement (*right*). The tertiary interaction between the substrate and the stem-loop V is represented by *shaded regions*. *Dotted lines* indicate the polarity of the phosphodiester backbone. The cleavage site is indicated by an *arrow*. The A730 loop is indicated, and the catalytically important nucleotide A756 is *circled*

bind between stems II and VI, and can be tethered to this position via a variety of manners. For instance, it can be attached to the 3′-end of stem II[103] or to the end of stem VI via a linker (D.A. Lafontaine and D.M.J. Lilley, unpublished results), both of which are readily accommodated within the model.

A number of experimental data suggest that the A730 loop is particularly important for the activity of the VS ribozyme. Indeed, using phosphorothioate incorporation, phosphate groups were identified (G757 and G758) in this loop that are important for cleavage activity.[102] Moreover, by using the "manganese rescue" variation, it was also found that the catalytic activity could be restored by using manganese ions, suggesting that metal ion binding sites are present in the loop that are important for the catalytic activity. In addition, the introduction of a variety of nucleotide analogues in this loop resulted in marked reduction in activity.[107]

Of all nucleotides of the A730 loop, the nucleotide A756 is the most sensitive position to point mutations, with most of the atomic positions of the nucleobase being important for activity, therefore suggesting a critical role for A756 in catalysis.[104,108–110] Supporting information has been obtained from cross-linking experiments.[111] Indeed, Collins and coworkers used a covalently attached substrate containing 4-thiouridine at the cleavage site and irradiated the sample with UV light. A cross-link to A756 was obtained, suggesting the close proximity of A756 to the cleavage site. Thus, these results strongly suggest that the A756 nucleobase is a key player during the catalytic reaction of the VS ribozyme. One characteristic aspect of nucleobase-driven catalysis is the dependence of cleavage activity on pH, which has been recently observed using ligation[110] and cleavage assays.[112] Additional evidence about the important role of A756 was obtained using NAIM analysis, which indicated the requirement for a protonated base at position 756.[107] In addition, the importance of A756 in the catalytic reaction has also been emphasized by covalently introducing an imidazole ring in place of the A756 nucleobase and by obtaining significant cleavage activity.[113] A similar procedure was used to show the importance of C75 in the context of the HDV ribozyme.[114] It is likely that a variety of strategies are in place in the VS ribozyme to achieve catalysis, and further experiments will have to be designed to fully grasp the catalytic reaction.

2.3 Riboswitches

Regulation of biological activity is an important cellular process for an organism to be viable. Well-known cellular controls are those involving proteins that regulate at the levels of transcription, translation, and mRNA stability.[115–123] *trans*-Acting RNA factors are also very important in the control of translation and stability of mRNAs,[122,124] as testified by microRNAs (miRNAs) and short-interfering RNAs (siRNAs); these are involved in a series of protein-mediated processing events that ultimately lead to the production of small RNA fragments which form base-paired complexes with target mRNAs, which are then digested by nuclease processing or by other mechanisms.[125,126] In most of these processes, RNA molecules are central

to the biological regulation, which testifies to the growing relevance of RNA in all areas of life.

Recently, a new type of noncoding RNA was uncovered that was found to control numerous fundamental genes using a novel regulation strategy. Indeed, riboswitches are RNA molecules carrying complex folded domains that are located in untranslated regions of mRNAs (recently discussed[127–130]). These molecules are able to sense the concentration of a target cellular metabolite, which is almost always related to the gene product encoded by the downstream sequence of the riboswitch. The sensing domain is called the aptamer and is highly conserved throughout evolution, whereas the expression platform, which is involved in genetic control, shows a much lower degree of conservation. The high degree of conservation of the aptamer most likely results from a selection pressure requiring the sensing of a cellular metabolite, which does not change through evolution. To modulate the expression of a given gene, the expression platform has the ability to fold into various secondary structures that are mutually exclusive. The genetic decision, which relies on the folding of the riboswitch, is dictated by the binding of the target metabolite to the riboswitch. In most cases, when the metabolite is in highly sufficient concentration, its binding to the riboswitch produces the shutdown of the synthesis of encoded biosynthetic proteins. To better illustrate this concept, one of the smallest riboswitches, the guanine-sensing riboswitch, is discussed briefly.

The guanine riboswitch is part of the purine riboswitch family and controls the expression of enzymes that are involved in the metabolism of purines. In *Bacillus subtilis*, it is present in the 5′-untranslated region (5′-UTR) of five transcriptional units, one of them being the *xpt-pbuX* operon that encodes a xanthine phosphoribosyltransferase and a xanthine-specific purine permease, respectively.[131] In the absence of guanine, the structure of the *xpt-pbuX* guanine riboswitch ("ON" state) is characterized by an antiterminator structure whose presence allows the transcription of the entire mRNA molecule and thus the expression of encoded proteins (Fig. 2.5). By contrast, upon binding of guanine to the aptamer domain, the

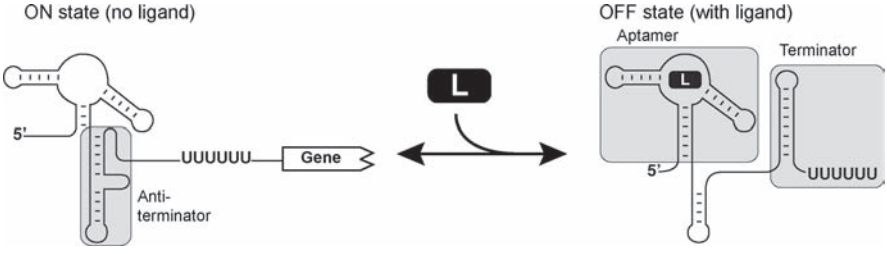

Fig. 2.5 Genetic control of the *xpt-pbuX* guanine riboswitch. In the absence of ligand, the guanine riboswitch adopts a secondary structure characterized by the presence of an anti-antiterminator stem, which allows the transcription of the complete mRNA to take place. However, upon ligand binding (uppercase *L*), the aptamer domain is stabilized, which in turn promotes the formation of a Rho-independent terminator structure, which is associated with the premature transcription termination

antiterminator structure is destabilized, allowing the folding of a Rho-independent terminator to take place, ultimately leading to the premature termination of transcription and to the inhibition of gene expression ("OFF" state). The regulation process is thought to occur without the aid of protein cofactors and thus provides a direct link between the metabolite being monitored and the regulated mRNA.

Various riboswitches have been discovered and shown to specifically recognize a variety of ligands. For instance, aptamer domains have been shown to bind adenine,[132] adenosylcobalamin,[133] flavin mononucleotide,[134,135] guanine,[9] glucosamine-6-phosphate,[123] glycine,[136] lysine,[137,138] S-adenosylmethionine (SAM),[139–142] and thiamine pyrophosphate (TPP).[134,143] The secondary structures of these aptamers are quite diverse (Fig. 2.6) and can be relatively simple (purine) or extremely complex (coenzyme B_{12}). Almost all aptamers shown in Fig. 2.6 control a single expression platform and exhibit a single sensing domain per expression platform. However, the glycine riboswitch (Fig. 2.6) is unique in the sense that it contains a dual aptamer which can sense two glycine molecules. Interestingly, Breaker and coworkers have shown that this riboswitch uses cooperativity to perform a tight regulation control in presence of glycine, and it has been suggested that this tandem configuration provides the riboswitch with a much narrow window toward glycine sensing and thus a higher selectivity in terms of gene activation and shutdown.[136] Moreover, it is important to realize that not all riboswitches negatively regulate expression upon ligand binding. The adenine and the glycine riboswitches are indeed performing a positive regulation in the presence of their respective ligand.[132,136] Riboswitches can also regulate at the level of translation, and in this case, it is not an intrinsic transcription terminator that is modulated but

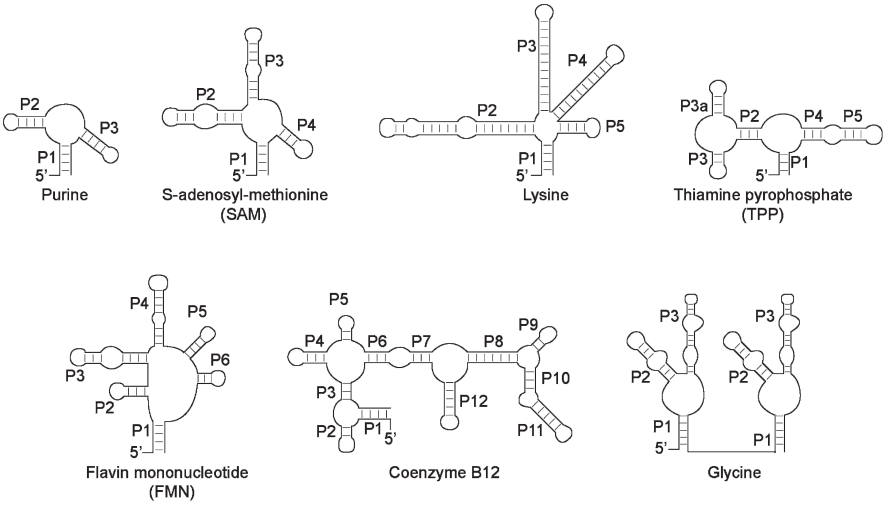

Fig. 2.6 Conventional riboswitches known to regulate on a single expression platform. Only the aptamer region is shown, and the monitored ligand is indicated for each one. The purine riboswitch comprises the adenine and the guanine riboswitches

a helical domain embedding Shine–Dalgarno and translation initiation sequences. Regardless of whether it is transcription or translation that is being regulated, modulation of the RNA structure is always the mechanistic step that controls gene expression, thus making RNA the key player in these regulatory processes.

A few new types of riboswitches have recently been discovered and are referred here as "atypical riboswitches" (Fig. 2.7). A dimeric structure embedding two common single-unit riboswitches, one responding to SAM and the other to coenzyme B_{12}, has recently been characterized in the *Bacillus clausii metE* mRNA.[144] This unique tandem configuration was shown to yield a gene control system functioning as a two-input NOR logic gate. This type of regulatory control expects that either of two chemical inputs yields an output of gene expression, but does not require that the two riboswitches influence each other. However, given that most riboswitches do not require such a tandem arrangement, why does the *B. clausii metE* gene need such a complex riboswitch configuration to be regulated? It turns out that the presence of two distinct riboswitches can be established by examining the metabolic pathway for methionine biosynthesis.[144] Indeed, in *B. clausii*, two proteins (MetE and MetH) are expressed that can independently use homocysteine to produce methionine, which is in turn used to make SAM. As a result, both *metE* and *metH* carry a SAM riboswitch as a control element in response to SAM. While MetE is able to perform the catalytic reaction, MetH uses the more reactive cofactor MeCbl, which is a derivative of AdoCbl. Thus, because cells can much more efficiently produce methionine by expressing MetH rather than MetE, it is energetically unfavorable to sustain MetE synthesis when there is plentiful AdoCbl.[144–146]

An additional atypical riboswitch was recently found in the context of the *B. subtilis glmS* gene.[123] The *glmS* gene encodes the enzyme glutamine-fructose-6-phosphate amidotransferase, which converts fructose-6-phosphate and glutamine into GlcN6P. The secondary structure of this riboswitch is shown in Fig. 2.11 (see later in this chapter). Strikingly, upon ligand binding to the RNA, an autocatalytic cleavage reaction takes place resulting in the scission of the mRNA molecule. The *glmS*

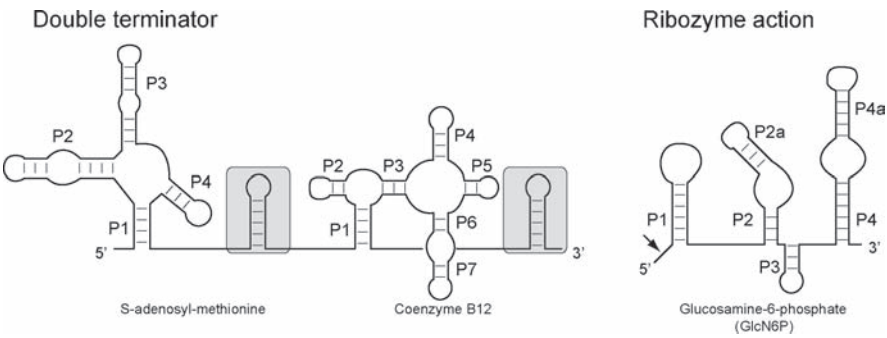

Fig. 2.7 Interestingly, several additional types of tandem configurations were recently reported by Breaker and coworkers,[144] suggesting that relatively simple RNA elements can be assembled to make sophisticated genetic decisions without involving protein factors

riboswitch, or the *glmS* ribozyme, is thus a ligand-dependent processing element that mixes both qualities of riboswitches and ribozymes to achieve genetic regulation of the synthesis of GlcN6P. The structural and mechanistic details of this regulatory RNA are discussed later in this chapter.

Given their natural propensity to bind small cellular metabolites and regulate essential genes, high expectations have been put on riboswitches as antimicrobial agents, partly based on previous successes using small-molecules targeting ribosomes.[147] By using riboswitches as drug targets, it is expected that artificial compounds could be used to bind riboswitches and inhibit the expression of the regulated cellular metabolites, thus heavily destabilizing bacterial pathogens. A proof-of-concept has already been done using the lysine riboswitch.[148] Indeed, several lysine analogues were identified to bind the riboswitch and inhibit cellular growth, most probably by inhibiting the expression of the gene regulated by the riboswitch. Surprisingly, Breaker and coworkers have determined that aminoethyl cysteine (AEC), originally characterized almost 50 years ago,[149] inhibits bacterial growth by targeting the lysine riboswitch.[138,148] Whether riboswitches will really serve as important drug targets needs to be determined, but nevertheless they provide a new avenue that is worthwhile to explore.

Several crystal structures have recently been solved for various riboswitches, which have revealed how ligand binding can be harnessed by RNA molecules to drive gene expression. The following sections are dedicated to discussing each of these in detail.

2.3.1 The Adenine Riboswitch

Purine riboswitches are among the smallest riboswitches that activate or inhibit gene expression in the presence of adenine or guanine, respectively (see Fig. 2.6). Despite the high structural similarity shown between the two aptamers, they display very high specificity, discrimination, and affinity toward their cognate ligand.[9,132] Although regulation by the guanine riboswitch appears to strictly modulate transcription, the adenine riboswitch regulates transcription and also translation. Given that the two riboswitches are very similar in structure, we focus here on the adenine-sensing riboswitch as it has been studied in more detail.

The adenine aptamer folds as a three-way junction where the P1 stem is the only helical region showing some degree of conservation (Fig. 2.8). Helices P2 and P3 are not conserved to a high degree but are involved in a loop–loop interaction.[150,151] The single-stranded core domain of the aptamer is the most conserved region and has been shown to undergo a structural reorganization upon ligand binding using in-line probing assays.[132] This assay exploits the inherent chemical instability of RNA under physiological conditions that is primarily the result of the cleavage of phosphodiester linkages.[152] Spontaneous scissions are more pronounced in unstructured single-stranded regions because internucleotide linkages are free to adopt an in-line conformation that is precluded in the context of an A-type helix.[152]