RNA Structure Determination

Methods and Protocols

Edited by

Douglas H. Turner

Department of Chemistry and Center for RNA Biology, University of Rochester
College of Arts and Sciences, Rochester, NY, USA

David H. Mathews

Department of Biochemistry and Biophysics and Center for RNA Biology,
University of Rochester Medical Center, Rochester, NY, USA
Research in the last 30 years has revealed that an unexpectedly large fraction of genomic DNA is transcribed into RNA [1]. Moreover, many new functions of RNA are being discovered [2]. This has provided a need for ways to rapidly translate sequence into structural information.

The twenty-first century witnessed many advances in modeling and determining RNA structures. Secondary structure prediction on the basis of sequences alone is increasingly accurate. New methods have been developed for experimentally probing secondary structure to identify paired and unpaired nucleotides for restraining predictions. Multiple methods are being developed to model three-dimensional structure. At the same time, more and more three-dimensional structures are being determined. The new structures are providing benchmarks for improving predictions of three-dimensional structure from sequence.

**Twenty-First Century Advances**

Secondary structure prediction improved in accuracy as a result of several innovations. New parameter sets were derived to quantify structure quality [3–5]. New algorithms were invented to consider folding of structural ensembles, rather than only most likely structures [4, 6–10]. Additionally, new approaches are available to determine the conserved secondary structure for multiple homologous sequences, thus increasing accuracy relative to single sequence structure prediction [11–17].

Probing structure by enzymatic and chemical methods is a cornerstone of determining RNA secondary structure [18–20]. New methods for probing structure were developed. In particular, a new class of chemical probing agents, based on selective 2′-hydroxyl acylation and primer extension readout (SHAPE), was developed to identify RNA nucleotides in flexible regions of the structure. The most reactive nucleotides tend to be in loops [21–23]. Unlike base-specific agents, SHAPE attacks 2′-hydroxyl groups and thus interrogates all nucleotides. SHAPE was also coupled with quantification of the reactivity per nucleotide, and these data provide restraints that dramatically improve the accuracy of secondary structure prediction [24–27]. At the same time, traditional probing agents were applied in new ways. Enzymatic cleavage was coupled with next-generation sequencing to probe structure across the transcriptome [28, 29]. The extent of dimethyl sulfate (DMS) reactivity was quantified and also used as restraints for structure prediction [30]. SHAPE reagents were shown to be effective at in vivo mapping [31, 32], as previously shown for DMS [33]. Finally, DMS and SHAPE were coupled with next-generation sequencing to probe RNA structure in vivo across the transcriptome [34–36].

Modeling of three-dimensional RNA structure has also advanced. As for protein structure prediction [37], RNA structure prediction uses blind modeling to assess advances in the field by employing new benchmarks, called RNA Puzzles [38, 39]. A number of groups participate in the blind predictions, using approaches ranging from physics-based to knowledge-based [40–49]. The second RNA Puzzles comparison concluded that overall
topologies are correctly modeled, but that noncanonical pair interactions are not yet well predicted \[39\].

At the start of the century, x-ray crystal structures of ribosomes were solved. Since that time, ongoing advances in x-ray crystallography \[50–52\], nuclear magnetic resonance (NMR) \[53–55\], and cryo-electron microscopy (cryo-EM) \[56–58\] have all led to the determination of more complex and higher resolution structures. Small angle x-ray scattering (SAXS) is being applied to RNA to determine molecular envelopes in solution \[59, 60\]. The advances extend to new approaches to consider ensembles and structural flexibility \[61, 62\]. Importantly, this work was enabled by development of new modeling methods, including improved methods for validating structures \[63\].

**Organization of the Book**

This book provides protocols for RNA structure modeling and determination. The first chapters provide protocols for RNA secondary structure prediction. Chapters 1 and 2 discuss single sequence modeling with the software packages, Crumple \[64\] and RNAstructure \[65\], respectively. Chapter 3 discusses using RNAstructure to model conserved secondary structures with multiple homologs. The prediction of bimolecular secondary structures with RNAstructure is presented in Chapter 4 and with Vfold \[42\] in Chapter 5. Chapter 6 presents STarMir \[66\], an application of secondary structure prediction to miRNA target prediction.

Chapters 7, 8, and 9 provide protocols for structure mapping, with traditional chemical agents \[18\], with enzymatic mapping across the transcriptome \[29\], and with SHAPE reagents \[67\], respectively. Chapter 10 provides protocols for using mapping data to constrain or restrain RNA secondary structure prediction with RNAstructure. Chapter 11 gives the protocol for using unassigned NMR resonances to improve secondary structure prediction and to provide initial assignments of some resonances to start solving a three-dimensional structure \[68\].

The book concludes with protocols focusing on three-dimensional structure. Chapters 12, 13, 14, and 15 provide modeling protocols for FARFAR \[43\], RNAComposer \[40\], ModeRNA \[41\], and MC-Fold \[46\], respectively. Chapter 16 provides an introduction to structure determination by NMR. Chapter 17 provides a protocol for x-ray crystallography determination of RNA structure.

**Rochester, NY, USA**

David H. Mathews

Rochester, NY, USA

Douglas H. Turner

**References**


# Contents

*Preface* ............................................................ v

*Contributors* ................................................... xiii

1 Crumple: An Efficient Tool to Explore Thoroughly the RNA Folding Landscape .................................................. 1
   *Ivan Guerra and Susan J. Schroeder*

2 Secondary Structure Prediction of Single Sequences Using RNAstructure .................................................. 15
   *Zhenjiang Zech Xu and David H. Mathews*

3 Prediction of Secondary Structures Conserved in Multiple RNA Sequences .................................................. 35
   *Zhenjiang Zech Xu and David H. Mathews*

4 Predicting RNA–RNA Interactions Using RNAstructure .......................................................... 51
   *Laura DiChiacchio and David H. Mathews*

5 A Method to Predict the Structure and Stability of RNA/RNA Complexes .................................................. 63
   *Xiaojun Xu and Shi-Jie Chen*

6 STarMir Tools for Prediction of microRNA Binding Sites .......................................................... 73
   *Shaveta Kanoria, William Rennie, Chaochun Liu, C. Steven Carmack, Jun Lu, and Ye Ding*

7 Traditional Chemical Mapping of RNA Structure in vitro and in vivo .................................................. 83
   *Pierre Fechter, Delphine Parmentier, Zongfu Wu, Olivier Fuchsbauer, Pascale Romby, and Stefano Marzi*

8 High-Throughput Nuclease Probing of RNA Structures Using FragSeq .................................................. 105
   *Andrew V. Uzilov and Jason G. Underwood*

9 Mapping RNA Structure in vitro with SHAPE Chemistry and Next-Generation Sequencing (SHAPE-Seq) .......................................................... 135
   *Kyle E. Watters and Julius B. Lucks*

10 Experiment-Assisted Secondary Structure Prediction with RNAstructure .................................................. 163
    *Zhenjiang Zech Xu and David H. Mathews*

11 RNA Secondary Structure Determination by NMR .......................................................... 177
    *Jonathan L. Chen, Stanislav Bellaousov, and Douglas H. Turner*

12 Modeling Small Noncanonical RNA Motifs with the Rosetta FARFAR Server .................................................. 187
    *Joseph D. Yesselman and Rhiju Das*

13 Automated RNA 3D Structure Prediction with RNAComposer .......................................................... 199
    *Marcin Biesiada, Katarzyna J. Purzycka, Marta Szachniuk, Jacek Blazewicz, and Ryszard W. Adamiak*

14 RNA 3D Structure Modeling by Combination of Template-Based Method ModeRNA, Template-Free Folding with SimRNA, and Refinement with QRNAs .......................................................... 217
    *Pawel Piatkowski, Joanna M. Kaspdzak, Deepak Kumar, Marcin Magnus, Grzegorz Chojnowski, and Janusz M. Bujnicki*
15  Exploring Alternative RNA Structure Sets Using
    MC-Flashfold and db2cm .................................................. 237
    Paul Dallaire and François Major

    Scott D. Kennedy

17  The Quick and the Dead: A Guide to Fast Phasing of Small Ribozyme
    and Riboswitch Crystal Structures ..................................... 265
    Jermaine L. Jenkins and Joseph E. Wedekind

Index ................................................................. 281
Contributors

Ryszard W. AdamiaK • Department of Structural Chemistry and Biology of Nucleic Acids, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland

Stanislav Bellaousov • Center for RNA Biology, University of Rochester, Rochester, NY, USA; Department of Biochemistry & Biophysics, University of Rochester Medical Center, Rochester, NY, USA

Marcin Biesiada • European Center for Bioinformatics and Genomics, Institute of Computing Science, Poznan University of Technology, Poznan, Poland; Department of Structural Chemistry and Biology of Nucleic Acids, Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland

Jacek Blazewicz • Institute of Computing Science, Poznan University of Technology, Poznan, Poland

Janusz M. Bujnicki • Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology, Warsaw, Poland; Laboratory of Bioinformatics, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland

C. Steven Carmack • Wadsworth Center, New York State Department of Health, Center for Medical Science, Albany, NY, USA

Jonathan L. Chen • Department of Chemistry, University of Rochester, Rochester, NY, USA; Center for RNA Biology, University of Rochester, Rochester, NY, USA

Shi-Jie Chen • Department of Physics, Informatics Institute, University of Missouri, Columbia, MI, USA; Department of Biochemistry, Informatics Institute, University of Missouri, Columbia, MI, USA

Grzegorz Chojnowski • Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology, Warsaw, Poland

Paul Dallaire • Department of Computer Science and Operations Research, Institute for Research in Immunology and Cancer, Université de Montréal, Montréal, QC, Canada

RhiJu Das • Biochemistry Department, Stanford University, Stanford, CA, USA

Laura DiChiaccio • Department of Biochemistry & Biophysics, University of Rochester Medical Center, Rochester, NY, USA; Center for RNA Biology, University of Rochester Medical Center, Rochester, NY, USA

Ye Ding • Wadsworth Center, New York State Department of Health, Center for Medical Science, Albany, NY, USA

Pierre Fechter • Biotechnologie et Signalisation Cellulaire, CNRS-INSERM, ESBS, Université de Strasbourg, Illkirch, France

Olivier FuchsbaUer • Architecture et Réactivité de l’ARN, CNRS, IBMC, Université de Strasbourg, Strasbourg, France

Ivan Guerra • Department of Chemistry & Biochemistry, Department of Microbiology & Plant Biology, University of Oklahoma, Norman, OK, USA

Jermaine L. Jenkins • Department of Biochemistry & Biophysics, University of Rochester School of Medicine and Dentistry, Rochester, NY, USA; Center for RNA Biology, University of Rochester School of Medicine and Dentistry, Rochester, NY, USA; The Structural Biology and Biophysics Facility, University of Rochester School of Medicine and Dentistry, Rochester, NY, USA
The Structural Biology and Biophysics Facility, University of Rochester School of Medicine and Dentistry, Rochester, NY, USA

ZONGFU WU • College of Veterinary Medicine, Nanjing Agricultural University, Nanjing, China

XIAOJUN ZECH XU • Department of Physics, Informatics Institute, University of Missouri, Columbia, MI, USA; Department of Biochemistry, Informatics Institute, University of Missouri, Columbia, MI, USA; Center for RNA Biology, University of Rochester Medical Center, Rochester, NY, USA

ZHENJIANG ZECH XU • Department of Biochemistry & Biophysics, University of Rochester Medical Center, Rochester, NY, USA; Center for RNA Biology, University of Rochester Medical Center, Rochester, NY, USA

JOSEPH D. YESSELMAN • Biochemistry Department, Stanford University, Stanford, CA, USA