

Rational Design of RiboNucleic Acids

Sven Findeiß^{*1}, Christoph Flamm^{*2}, and Yann Ponty^{*3}

1 Leipzig University, DE. sven@bioinf.uni-leipzig.de

2 University of Vienna, AT. xtof@tbi.univie.ac.at

3 Ecole Polytechnique - Palaiseau, FR. yann.ponty@lix.polytechnique.fr

Abstract

This report documents the program and outcomes of Dagstuhl Seminar 22381 “Rational Design of RiboNucleic Acids” (RNAs). The seminar covered a wide array of models, algorithmic strategies, molecular scales and modalities, all targeting *in silico design* of RNAs performing predefined biological functions. It consisted in a series of talks, each being allocated a generous time budget enabling frequent (welcomed!) interruptions and fruitful discussions. Applications of rational RNA design include mRNA vaccines; RNAs acting as sensors; self-replicating RNAs, relevant to RNA world/origin of life studies; populations of RNAs performing computations, e.g. through strand-displacement systems; RNA origamis forming nano-architectures through self-assembly; weakly interacting RNAs inducing the formation of droplets within cells through liquid-liquid phase separation. Those diverse applications are typically tackled by Bioinformatics-inclined scientists, contributing to distinct areas of life science and, as a result, somewhat isolated and sometimes unaware of similar pursuits in neighboring fields. The overarching goals of this meeting were to gather computational scientists from multiple fields, increase awareness of relevant efforts in distant communities, and ultimately contribute to a transversal perspective where RNA design becomes an object of study in itself.

Seminar September 18–23, 2022 – <http://www.dagstuhl.de/22381>

2012 ACM Subject Classification Applied computing → Bioinformatics; Applied computing → Molecular evolution; Applied computing → Molecular structural biology; Theory of computation → Discrete optimization; Theory of computation → Dynamic programming; Theory of computation → Parameterized complexity and exact algorithms; Mathematics of computing → Optimization with randomized search heuristics; Computing methodologies → Discrete space search; Computing methodologies → Reinforcement learning; Computing methodologies → Neural networks

Keywords and phrases RNA, RNA design, Inverse folding, RNA structure, mRNA design, RNA sensors, Co-transcriptional folding, Molecular evolution, Distant homology, Drug design

Digital Object Identifier 10.4230/DagRep.12.9.121

1 Executive Summary

Sven Findeiß

Christoph Flamm

Yann Ponty

License © Creative Commons BY 4.0 International license
© Sven Findeiß, Christoph Flamm, and Yann Ponty

Context and selected takeaways

RiboNucleic Acids (RNAs) are ubiquitous macromolecules within biological systems, capable of performing a wide range of regulatory and catalytic functions. This versatility can

* Editor / Organizer



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 4.0 International license

Rational Design of RiboNucleic Acids, *Dagstuhl Reports*, Vol. 12, Issue 9, pp. 121–149

Editors: Sven Findeiß, Christoph Flamm, and Yann Ponty



Dagstuhl Reports

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

be harnessed, and RNAs are increasingly utilized to accurately monitor and control biological processes [19], leading to RNA being found at the core of modern therapeutics [18]. It is therefore not surprising that the RNA-guided CRISPR-Cas9 editing [10], rewarded by the 2020 Nobel Prize in Chemistry, and mRNA-based vaccines [12], are at the forefront of modern biotechnology. For many functional RNA families [11], decades of research have produced a deep understanding of the sequence and structural basis underlying their biological function(s). Such studies, coupled with mature computational methods for structure prediction [23], have paved the way for a rational design of RNAs targeting a wide diversity of biological function [8, 2, 13].

Accordingly, RNA design has emerged as an exciting open computational problems in molecular biology. Owing to the discrete nature of RNA sequence and popular structural representations (e.g. secondary structure), RNA design has inspired the contribution of a large number of diverse algorithms [9, 20, 14, 4] for the inverse folding problem, i.e. the design of an RNA sequence which preferentially and effectively folds into a predefined (secondary) structure. Given the, recently established, NP-Hardness of the problem, even for minimal energy models [1], many of those algorithmic predictions are either heuristics, exponential-time or based on a variety of machine learning techniques.

More generally, RNA design addresses the generation of sequences of nucleotides targeting a given biological function. A non-exhaustive list of classic design objectives includes:

- Preferential adoption of one or multiple given structures (inverse folding);
- Sequence specific constraints such as an overall (di-)nucleotide composition [21], encoding of an amino-acid sequence (mRNA design), presence/absence of motifs [22];
- Adoption of different conformations upon presence of ligand (RNA switches and sensors) [3];
- Effective and specific interactions with targeted partners (RNAs, proteins) cascading into system-level regulatory effects [15, 16];
- Self-assembly into large scale architectures, ultimately adopting a predefined 3D shape (RNA origami) [6];
- Exploit co-transcriptional folding, and more general out-of-equilibrium regimes to perform computations (strand displacement systems, oritatami) [5]

Typical applications of design include novel therapeutic strategies, control principles for existing biological systems, or sensors for the presence of small molecules [3], but designed sequences can also provide an objective experimental assessment of functional hypotheses, where designs are synthesized and their effect on the cellular context can be tested *in vitro* and, in turn, *in vivo*.

Over the course of the seminar, we witnessed a substantial recent expansion of the scope of applications. Beyond classic but still challenging objectives of design, including riboswitches addressed by Talk 5.8, 5.9 and 5.21, messenger RNAs towards vaccine objectives mentioned by Talk 5.27, and CRISPR gRNAs mentioned by Talk 5.11, novel applications of RNA design emerged during the seminar. Talk 5.11 introduced SARS CoV 2 sensors based on strand displacement, Talk 5.15 addressed self-replicating ribozymes connected with origin-of-life questions, and Talk 5.6 explored rational design principles for repetitive RNAs inducing the formation of cellular droplets through liquid-liquid phase separation.

RNA Design as a discrete (inverse) optimization problem

The inverse folding problem, one of the central elements of RNA design, is a hard computational problem [1]. Although attracting a wide interest from the community, it is also one of the very few problems in computational biology whose complexity status has remained

open for a long time (about three decades). This difficulty can be attributed to a lack of a suitable conceptual framework for inverse combinatorial problems. Indeed, inverse folding can be viewed as the search of a pre-image, in a function that maps each RNA sequence to its most stable conformation, the latter being computed using a polynomial, yet non-trivial, dynamic programming algorithm [23]. Natural generalizations virtually include any instance of inverse optimization problems, and could be of general interest to the Computer Science research field. Prior works in this direction have led to characterization of designable structures based on formal languages and graph theory [7], revealing strong connections to many subfields of computer science (for instance, between positive design and graph coloring).

In Talk 5.18 it was discussed that a flexible inverse folding approach, e.g. by allowing the extension of helices by at most one base pair, seems to be easier than keeping the problem strict. Such a flexibility in the structural objective of design was also emphasized as desirable by Talk 5.5. The problem of classical inverse folding can be extended from one to multiple target structures, and Talk 5.27 showed that this can be solved by an elegant dynamic programming approach that is fixed parameter tractable. The resulting framework was further generalized, and is not only applicable to RNA design, but also to apparently more distant problems such as the alignment of RNAs with pseudoknots. *In silico* designs and analysis depend on the accuracy of the applied energy model. In Talk 5.12 it has been underpinned that a systematic perturbation of parameters can be used to define a notion of robustness of individual parameters of an energy model, and help to improve prediction accuracy. Talk 5.28 revisited the inverse folding problem as an inverse optimization problem, and showed that many local structural motifs do not admit a design, with consequences to the space of designable structures, but raised fundamental questions on a relatively new flavor of optimization.

RNA Design in Structural Bioinformatics

Inverse folding also represents the ultimate test of our understanding of the mechanisms governing the folding of macromolecules. Given a set of folding rules (typically, an energy model), a synthesis of *in silico* designed sequences combined with high-throughput experiments (e.g., structure probing) enables an assessment of the compatibility of the determined structure with the initial target. Observed discrepancies can then be used to assess the quality of predictive models, especially those based on statistical potentials which may be prone to overfitting. Systematic local imprecisions can also be used to refine energy models, enabling the generation of better designs, whose iteration represents a virtuous circle, ultimately contributing to a better understanding of folding principles.

A nascent RNA molecule typically folds during its transcription. Frameworks to simulate this kinetically driven process can help to interpret experimental results (Talk 5.4) but as neither the simulation nor the experiment is perfect, quality assurance (Talk 5.14) of the *in silico* investigations is essential and results have to be interpreted with caution, as for instance the mapping of time scales is a non trivial task. Finally, complex RNA hybridization networks are designed *in silico* to perform regulatory functions with complex temporal dynamics. A simplified kinetic model, introduced in Talk 5.26, for RNA/RNA hybridization represents an attractive evaluation model for the design of interactions.

At a much more detailed 3D level, Talk 5.17 showed that high-resolution experimental techniques can be used to observe dynamic behaviors, sometimes triggered by the binding of a ligand, and could inform future objective functions. Talk 5.24 and 5.16 described coarse-grained models amenable to molecular dynamics. Interestingly, the latter can be leveraged in order to study kinetics behaviors at the 3D level.

RNA Design in Synthetic Biology and Natural Computing

This line of research applies various engineering principles to the design, and construction of artificial biological devices. While initially focused on hijacking naturally-occurring regulatory functions through a copy/paste of evolved genetic material [17], the need for a precise control and for a modularity/orthogonality of constructs, has increasingly led to a *de novo* design based on nucleic acids. Recently, RNA has been successfully used as a material for the design of whole regulatory circuits, or for the construction of complex programmable shapes (RNA origamis [6]), with promising applications as biomaterial.

Software frameworks, like the ones presented in Talk 5.10 and Talk 5.22, make the construction of large DNA and RNA nano-structures possible. Those designs are not only adopting the right structure *in silico* according to combinatorial folding algorithms, but can also be validated by simulations (Talk 5.24) and microscopy (Talk 5.1). This observation suggests that the difficulty of design could stem from the compactness of targeted ncRNAs, while larger (but more regular) RNAs may be easier to design, an element that could inspire future theoretical studies.

In the course of the seminar it became evident that information from the 2D and 3D level need to be mapped onto each other. Design could therefore benefit from multiscale approaches: selecting candidates with 2D objectives, use coarse grained 3D analysis (Talk 5.16) and go to a full atom final validation for critical sub-regions. The curation of refined and non-redundant 3D RNA structures (Talk 5.2) and the systematic extraction of information from such a data set can help to investigate for instance structural features of modified bases or to propose isosteric structural mutations (Talk 5.19) in order to generalize the design from 2D to compact 3D architectures.

Programmable RNA folding can also be used as a computational model, allowing for the computation of complex programs based on cotranscriptional folding phenomena. RNA regulatory circuits can be used to emulate Boolean functions, allowing a precise and expressive control of regulatory networks at an early stage of the gene expression process. Talk 5.23 introduced RNA oritatami, a Turing-complete model of computation based on cotranscriptional folding inspired by cellular automata. Talk 5.7 described exciting applications of design to generate easily-checkable QR codes that reveal contamination in a closed environment. However an application-agnostic implementation of the strand displacement systems underlying some of those applications still represent major challenges in RNA, as shown and discussed during Talk 5.3. Those include intra-molecular base pairs and an overall wasteful behavior that motivates efficient recycling strategies.

RNA evolution and Machine Learning

The analysis of new RNA families, such as the pervasive and poorly understood lncRNAs or the numerous viral/bacterial non-coding RNAs observed in metagenomics experiments, relies critically on the identification of an evolutionary pressure, allowing to hypothesize new functions. Given a family of homologous RNAs sharing established functional traits, it is classic to ask whether an observed property, such as the occurrence of a common motif or a given covariation pattern, is likely to reveal a yet-unknown selective pressure or, conversely, is merely the consequence of established functional traits. Classic bioinformatics methods rephrase the problem in a hypothesis-testing framework, and compute the probability that a sequence, generated at random in a model that captures existing constraints, features the observed property. Ideally, such sequences should represent solutions to an instance of the design problem, target established functions, while respecting a distribution that can either be derived from the targeted function, or learned from data.

Talk 5.5 presented a context where rational design methodologies were utilized to capture remote homologs of a suspected, but scarcely-populated, functional family of ncRNAs. Generative models can also be used for design, in cases where the underlying model of function is partially understood, and should be learned from the data. Talk 5.8 used Restricted Boltzmann Machines (RBM), an unsupervised learning approach, to pick up the intricate probability distribution of the statistical features of a naturally-occurring riboswitch. The RBM was then used to generate novel instances with the same distribution of features, resulting in an enrichment of functional designs as revealed by experimental validation. Direct Coupling Analysis was also used in Talk 5.15 to generate self-replicating ribozymes, using a complex definition of function that may require some element of learning. Interestingly, the efficacy of designs was ultimately shown to benefit from further refinements using classic combinatorial methods for inverse folding, suggesting future hybrid ML/combinatorial methodologies.

While RNA design is an increasingly important computational task in molecular biology, nanotechnology and medicine, methods for computational rational design are still lacking for many applications. Moreover, many design tasks are currently addressed using algorithmic techniques (e.g. Markov chain Monte Carlo) that are clearly superseded by the state-of-the-art in algorithmic research. Conversely, computer scientists considering design tasks usually limit themselves to inverse folding, overlooking a rich bestiary of computational problems whose consideration would, in turn, undoubtedly lead to the emergence of new algorithmic paradigms.

Talk 5.20 showed that basic ML architectures can be learned in the context of reinforcement learning and can be successful for basic inverse folding of RNA. Talk 5.25 presented a complete design story, describing a methodology to advance our understanding of tRNAs. In particular, a mechanistic understanding of the target function can be gained by masking constraints during redesign, and the differentiability of the design problem can lead to great speedups of the computation. However, ML approaches may not always represent a silver bullet in RNA bioinformatics, and Talk 5.13 dramatically illustrated this in the context of RNA folding, a context where the quality and biases in the data strongly impacts, and probably hinders for intrinsic reasons, the predictive capabilities of deep learning-based methods. As a consequence, synthetic data can and should be used to test the capacity of learning architectures on simplified problems before embarking into “real life” learning.

References

- 1 Édouard Bonnet, Paweł Rzżewski, and Florian Sikora. Designing RNA secondary structures is hard. In Benjamin J. Raphael, editor, *Research in Computational Molecular Biology - 22nd Annual International Conference, RECOMB 2018*, volume 10812 of *Lecture Notes in Computer Science*, pages 248–250, Paris, 2018. Springer.
- 2 Matthew G. Costales, Jessica L. Childs-Disney, Hafeez S. Haniff, and Matthew D. Disney. How we think about targeting RNA with small molecules. *Journal of Medicinal Chemistry*, 63(17):8880–8900, 2020. PMID: 32212706.
- 3 Sven Findeiß, Maja Etzel, Sebastian Will, Mario Mörl, and Peter F Stadler. Design of artificial riboswitches as biosensors. *Sensors (Basel, Switzerland)*, 17(9):E1990, August 2017.
- 4 Juan Antonio Garcia-Martin, Peter Clote, and Ivan Dotu. RNAiFOLD: a constraint programming algorithm for RNA inverse folding and molecular design. *Journal of Bioinformatics and Computational Biology*, 11(2):1350001, 2013.

- 5 Cody Geary, Pierre-Étienne Meunier, Nicolas Schabanel, and Shinnosuke Seki. Oritatami: A Computational Model for Molecular Co-Transcriptional Folding. *International Journal of Molecular Sciences*, 20(9):2259, May 2019.
- 6 Cody Geary, Paul W. K. Rothmund, and Ebbe S. Andersen. A single-stranded architecture for cotranscriptional folding of RNA nanostructures. *Science*, 345(6198):799–804, 2014.
- 7 Jozef Haleš, Alice Héliou, Ján Maňuch, Yann Ponty, and Ladislav Stacho. Combinatorial RNA Design: Designability and Structure-Approximating Algorithm in Watson-Crick and Nussinov-Jacobson Energy Models. *Algorithmica*, 79(3):835–856, Nov 2017.
- 8 Stefan Hammer, Christian Günzel, Mario Mörl, and Sven Findeiß. Evolving methods for rational de novo design of functional RNA molecules. *Methods*, 161:54–63, may 2019.
- 9 I.L. Hofacker, W. Fontana, P.F. Stadler, L.S. Bonhoeffer, M. Tacker, and P. Schuster. Fast folding and comparison of RNA secondary structures. *Monatsch. Chem.*, 125:167–188, 1994.
- 10 Martin Jinek, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna, and Emmanuelle Charpentier. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, 337(6096):816–821, 2012.
- 11 Ioanna Kalvari, Joanna Argasinska, Natalia Quinones-Olvera, Eric P Nawrocki, Elena Rivas, Sean R Eddy, Alex Bateman, Robert D Finn, and Anton I Petrov. Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic acids research*, 46:D335–D342, January 2018.
- 12 David M. Mauger, B. Joseph Cabral, Vladimir Presnyak, Stephen V. Su, David W. Reid, Brooke Goodman, Kristian Link, Nikhil Khatwani, John Reynders, Melissa J. Moore, and Iain J. McFadyen. mRNA structure regulates protein expression through changes in functional half-life. *Proceedings of the National Academy of Sciences*, 116(48):24075–24083, 2019.
- 13 Na Qu, Yachen Ying, Jinshan Qin, and Antony K Chen. Rational design of self-assembled RNA nanostructures for HIV-1 virus assembly blockade. *Nucleic Acids Research*, 50(8):e44–e44, 12 2021.
- 14 Vladimir Reinharz, Yann Ponty, and Jérôme Waldspühl. A weighted sampling algorithm for the design of RNA sequences with targeted secondary structure and nucleotide distribution. *Bioinformatics*, 29(13):i308–i315, 2013.
- 15 Guillermo Rodrigo, Thomas E Landrain, and Alfonso Jaramillo. De novo automated design of small RNA circuits for engineering synthetic riboregulation in living cells. *Proceedings of the National Academy of Sciences U S A*, 109(38):15271–6, 2012.
- 16 Guillermo Rodrigo, Thomas E. Landrain, Eszter Majer, José-Antonio Daròs, and Alfonso Jaramillo. Full design automation of multi-state RNA devices to program gene expression using energy-based optimization. *PLoS Computational Biology*, 9(8):e1003172, 2013.
- 17 E Westhof, B Masquida, and L Jaeger. RNA tectonics: towards RNA design. *Folding & design*, 1(4):R78–88, 1996.
- 18 Melanie Winkle, Sherien M. El-Daly, Muller Fabbri, and George A. Calin. Noncoding RNA therapeutics – challenges and potential solutions. *Nat Rev Drug Discov*, 20:629–651, 2021.
- 19 Sherry Y. Wu, Gabriel Lopez-Berestein, George A. Calin, and Anil K. Sood. RNAi therapies: Drugging the undruggable. *Science Translational Medicine*, 6(240):240ps7, 2014.
- 20 Joseph N Zadeh, Brian R Wolfe, and Niles A Pierce. Nucleic acid sequence design via efficient ensemble defect optimization. *Journal of Computational Chemistry*, 32(3):439–52, 2011.
- 21 Yang Zhang, Yann Ponty, Mathieu Blanchette, Eric Lecuyer, and Jérôme Waldspühl. SPARCS: a web server to analyze (un)structured regions in coding RNA sequences. *Nucleic Acids Research*, 41(Web Server issue):W480–5, July 2013.

- 22 Yu Zhou, Yann Ponty, Stéphane Vialette, Jérôme Waldispühl, Yi Zhang, and Alain Denise. Flexible RNA design under structure and sequence constraints using formal languages. In Jing Gao, editor, *ACM Conference on Bioinformatics, Computational Biology and Biomedical Informatics. ACM-BCB 2013, Washington, DC, USA, September 22-25, 2013*, page 229. ACM, 2013.
- 23 M. Zuker and P. Stiegler. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Research*, 9:133–148, 1981.

2 Table of Contents

Executive Summary	
<i>Sven Findeiß, Christoph Flamm, and Yann Ponty</i>	121
Participants and Group Composition	130
Overall Organization and Schedule	131
Overview of Talks	132
Structural basis of RNA origami design, folding and flexibility	
<i>Ebbe Sloth Andersen</i>	132
Datasets for benchmarking RNA design algorithms	
<i>Maciej Antczak</i>	133
On the compilation of multi-stranded nucleic acids circuits	
<i>Stefan Badelt</i>	133
Simulations of cotranscriptional folding explain the impact of sequence mutations	
<i>Stefan Badelt</i>	134
Eukaryotic riboswitch detection using inverse RNA folding	
<i>Danny Barash</i>	134
Design of RNA tandem repeats creating RNA droplets forming liquid-liquid phase separation	
<i>Sarah Berkemer</i>	134
Scaling and Limits of DNA Strand Displacement Computing	
<i>Harold Fellermann</i>	135
Generative modelling of riboswitches with restricted Boltzmann machines	
<i>Jorge Fernández de Cossío Díaz</i>	136
Get away from Plug and Pray: Synthetic Riboswitches - Applications and open Problems	
<i>Sven Findeiß</i>	136
Designing RNA during the DNA Origami revolution	
<i>Cody Geary</i>	137
Two design stories: probes for SARS-CoV-2 detection and CRISPR/Cas9 gRNAs	
<i>Jan Gorodkin</i>	137
What can geometric combinatorics say about RNA design?	
<i>Christine Heitsch</i>	138
Experiments in Deep Learning for RNA Secondary Structure Prediction	
<i>Ivo Hofacker</i>	138
BarMap-QA – Cotranscriptional folding with quality assurance	
<i>Felix Kühnl</i>	139
Discovering RNA Self-Reproducers By In Silico And In Vitro Screening	
<i>Philippe Nghe</i>	139
Physical modeling of RNA polymorphism	
<i>Samuela Pasquali</i>	140

RNA dynamics: one basepair at a time <i>Katja Petzold</i>	141
Minimalistic RNA inverse folding <i>Yann Ponty</i>	141
Challenges in designing RNA non-canonical modules <i>Vladimir Reinhartz</i>	142
Learning to Design RNA <i>Frederic Runge</i>	142
Differential SHAPE probing to screen computationally designed RNA and to detect pseudoknot and non-canonical interactions <i>Bruno Sargueil</i>	143
ENSnano: a 3D modeling software for designing complex DNA/RNA nanostruc- tures <i>Nicolas Schabanel</i>	143
Single-Stranded Architectures for RNA Co-Transcriptional Folding <i>Shinnosuke Seki</i>	145
Coarse-grained modeling for RNA nanotechnology <i>Petr Sulc</i>	145
Persuading tRNA to jump over stop codons <i>Andrew Torda</i>	146
Kinetic features of RNA-RNA interactions <i>Maria Waldl</i>	146
Infrared: A sampling framework for RNA design... and beyond <i>Sebastian Will</i>	147
Forbidden RNA motifs and the cardinality of secondary structure space <i>Hua-Ting Yao</i>	148
Participants	149

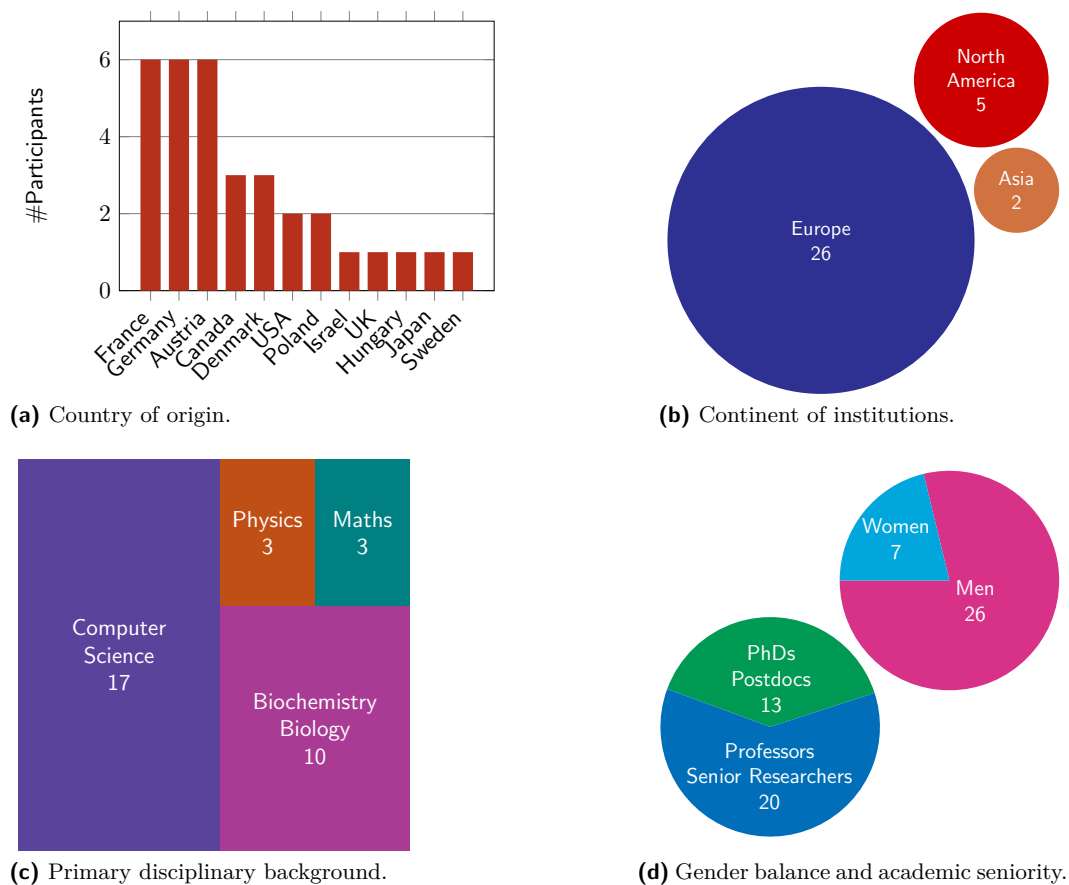


Figure 1 Participants statistics. Breakdown of workshop participants by country/continent of employing institution, primary disciplinary background, gender and academic seniority.

3 Participants and Group Composition

The list of participants included 33 researchers, selected for their outstanding contributions to RNA design, and/or their outstanding potential to impact future developments of the field. It should be noted that, in addition to those ultimate participants, a dozen confirmed researchers had to cancel, partly due to the ongoing pandemics.

As shown in Figures 1a and 1b, participants primarily originated from European institutions (26/33), but also from North America (5/33) and Asia (2/33), with five countries (Austria, France, Germany, Canada and Denmark) representing almost three quarters (24/33) of participants. While this concentration largely reflects the main centers of research on RNA bioinformatics, combined with the European location of the seminar center, the organizers regretted the absence of key players from North America and Asia, and will take this fact into consideration upon (possibly) organizing future editions of the seminar.

A key aspect of RNA design is that it requires a constant interdisciplinary dialogue, involving partners originating from diverse fields of research. Those include computational scientists to design algorithms and methods, modelers and experts in biochemistry to formulate models that are both accurate and computationally tractable, and end users/stakeholders from the fields of biology, biotechnology and medicine to assess the suitability of current

design models and inform future developments. The organizers are proud to report that the seminar was able to strike such a critical interdisciplinary mix, as shown in Figure 1c. While roughly half (17/33) of our participants were originally trained in Computer Science, one third (10/33) had initial background in Biology or Biochemistry, with the remaining participants being equally split between Physics and Mathematics (3/33 each). Interdisciplinary research also notoriously benefits from an early exposure of junior scientists to other fields of research. As shown in Figure 1d, the organizers were pleased to witness the presence of two fifths (13/33) of junior scientists (PhD candidates or Postdocs), among a majority (20/33) of more established scientists (Asst/Full prof. or permanent researchers).

Finally, partly due to late cancellation and despite a conscious mitigation effort by the organizers, the gender representation among participants certainly showed imbalance, with only a fifth of female researchers (7/33), Figure 1d. One possible reason, mentioned by some prospective participants, is the lack of support for daycare options (esp. towards small infants), which we understand is being considered by the center. While the organizers believe that this aspect is only partly within their control, it will nevertheless be the object of an increased focus while organizing future seminars.

4 Overall Organization and Schedule

Firstly, we wish to stress the impact of the ongoing COVID 19 pandemics on the organization of this seminar. Beyond the above-mentioned cancellations, the seminar was originally envisioned in February 2019, submitted in April 2019 and accepted in July of 2019, to be held in October 2020 and finally canceled due to a deep resurgence of COVID in Europe. A proposal was resubmitted in November 2020, and accepted in February 2021, to be finally held in September 2022. Overall, this workshop has been 3 1/2 years in the making, and the organizers were particularly excited to see it finally happen after such uncertainty.

The seminar itself consisted of talks, mostly scheduled before the seminar, while leaving ample time to impromptu discussions and spontaneous talk propositions. Wednesday afternoon was intentionally left open, to allow participants to interact in a less formal/public environment. We also finished the seminar after lunch on Friday, to allow most participants to reach home before the weekend. This left us with sufficient time to feature 28 talks, each of a duration of 30 or 45 minutes, structured in 7 sessions.

The seminar started with a joint talk by the organizers on Monday morning, aiming at providing sufficient context for all participants to follow and maximally benefit from the remaining talks. The afternoon session, named *Design Stories*, consisted of success stories in RNA design, a topic which we reasoned would expose most participants to the diversity of objectives required by applications, and base our future discussions on realistic use-cases. Tuesday morning's *Molecular Biology* session mentioned topics at the interplay of molecular modeling and evolution, while the afternoon's *Molecular Computing* session showcased design challenges and solutions arising in a context where RNA is used as a programmable material, capable of self-assembly and computation. The sole Wednesday morning session was dedicated to *Machine Learning* in the context of design, with a strong emphasis on generative models being used as a substitute for the classic specification/implementation philosophy. On Thursday, the morning *Combinatorial Design* session was dedicated to algorithmic and enumerative considerations of the yet-unsolved inverse folding problem, while the afternoon session focused on *3D and Design*, a very challenging context whose objective functions are still to be defined. Further increasing in difficulty, the Friday morning *Design for Dynamic Landscapes* session closed the seminar with contributions towards the design of RNA folding or interacting kinetically, out of the equilibrium regime.

Due to the interdisciplinary nature of the audience and topic, we ended up not eliciting to formally include an *open problem* session, although some formalized problems were mentioned during talks, and more thoroughly explored in smaller groups, past dinnertime.

5 Overview of Talks

5.1 Structural basis of RNA origami design, folding and flexibility

Ebbe Sloth Andersen (Aarhus University, DK)

License © Creative Commons BY 4.0 International license
© Ebbe Sloth Andersen

Joint work of Ebbe Sloth Andersen, Helena Rasmussen, Ewan K. S. McRae, Jianfang Liu; Andreas Bøggild, Michael Truong-Giang Nguyen, Néstor Sampredo Vallina, Thomas Boesen, Jan Skov Pedersen, Gnag Ren, Cody Geary

The research field of RNA nanotechnology develops methods for the rational design of self-assembling RNA nanostructures with applications in nanomedicine and synthetic biology. Inspired by the cotranscriptional folding of biological RNA molecules, we developed the RNA origami method to design RNA nanostructures compatible with cotranscriptional folding [1, 2], advantageous for large-scale production *in vitro* and expression *in vivo*. However, advancing this technology further will require a better understanding of RNA structural properties and the non-equilibrium dynamics of the cotranscriptional folding process. Here, we use cryogenic electron microscopy to study a panel of RNA origami structures at sub-nanometer resolution revealing structural parameters of kissing loop and crossover motifs, that are further used to optimize designs by reduction of internal strain and global twist. In three-dimensional bundle designs, we discover a novel kinetic folding trap that forms during cotranscriptional folding and is only released 10-12 hours after transcription start. We characterize the conformational landscape of RNA origamis to reveal the RNA flexibility of helices and structural motifs. Finally, we demonstrate that large distinctive RNA origami shapes are visible by cryo-electron tomography pointing to potential use as markers in cellular environments. Our results improve understanding of RNA structure, folding, and dynamics, providing a basis for rational design of genetically encoded RNA nanodevices.

References

- 1 Geary, C., Rothmund, P. W. and Andersen, E. S. A single-stranded architecture for cotranscriptional folding of RNA nanostructures. *Science* 345, 799-804, doi:10.1126/science.1253920 (2014).
- 2 Geary, C., Grossi, G., McRae, E. K. S., Rothmund, P. W. K. and Andersen, E. S. RNA origami design tools enable cotranscriptional folding of kilobase-sized nanoscaffolds. *Nat Chem*, doi:10.1038/s41557-021-00679-1 (2021).

5.2 Datasets for benchmarking RNA design algorithms

Maciej Antczak (Poznan University of Technology, PL)

- License** © Creative Commons BY 4.0 International license
© Maciej Antczak
- Joint work of** Maciej Antczak, Marta Szachniuk
- Main reference** Bartosz Adamczyk, Maciej Antczak, Marta Szachniuk: “RNAsolo: a repository of cleaned PDB-derived RNA 3D structures”, *Bioinform.*, Vol. 38(14), pp. 3668–3670, 2022.
URL <https://doi.org/10.1093/bioinformatics/btac386>
- Main reference** Jakub Wiedemann, Jacek Kaczor, Maciej Milostan, Tomasz Zok, Jacek Blazewicz, Marta Szachniuk, Maciej Antczak: “RNAloops: a database of RNA multiloops”, *Bioinform.*, Vol. 38(17), pp. 4200–4205, 2022.
URL <https://doi.org/10.1093/bioinformatics/btac484>

In this talk, we will present the databases developed to support benchmarking of bioinformatics algorithms targeting RNA, including the ones for RNA design. RNAsolo¹ collects experimentally determined 3D RNA structures from RNAs alone, protein-RNA complexes, and DNA-RNA hybrids and organizes them into classes of equivalent structures. Their sequences and tertiary structures are grouped in 192 benchmark sets ready for download and automated processing. RNAloops² aims to facilitate the study of multiloops in RNA molecules. It collects n-way junctions found in experimental RNA structures and allows to search them by sequence, secondary structure topology, or structure parameters. Both data sources address RNA-related studies by providing reliable sequence and structure data and efficient search facilities.

5.3 On the compilation of multi-stranded nucleic acids circuits

Stefan Badelt (Universität Wien, AT)

- License** © Creative Commons BY 4.0 International license
© Stefan Badelt
- Main reference** Stefan Badelt, Seung Woo Shin, Robert F. Johnson, Qing Dong, Chris Thachuk, Erik Winfree: “A General-Purpose CRN-to-DSD Compiler with Formal Verification, Optimization, and Simulation Capabilities”, in *Proc. of the DNA Computing and Molecular Programming - 23rd International Conference, DNA 23, Austin, TX, USA, September 24–28, 2017, Proceedings, Lecture Notes in Computer Science*, Vol. 10467, pp. 232–248, Springer, 2017.
URL https://doi.org/10.1007/978-3-319-66799-7_15

Compilation from high level languages to low level languages is a fundamental concept in computer science, and it enables researchers to program silicon-based machines even though they have no understanding of assembler code or transistors. In previous work, we have shown that a description of nucleic acid circuits at the domain level (equipped with an approximate biophysical model for DNA), can be formally derived from a high level language, e.g. compilation from a boolean circuit to formal chemical reaction network to a domain-level strand displacement system.

The next challenge is to ensure a correct compilation from the domain-level system specification to the nucleotide level, which must involve both nucleic acid sequence design and a verification based on folding kinetics at the secondary structure level. Currently, we are exploring new sequence design techniques for large and complex nucleic acid reaction networks, that also systematically incorporate feedback from experimental work.

¹ <https://rnasolo.cs.put.poznan.pl/>

² <https://rnaloops.cs.put.poznan.pl/>

5.4 Simulations of cotranscriptional folding explain the impact of sequence mutations

Stefan Badelt (Universität Wien, AT)

License © Creative Commons BY 4.0 International license
© Stefan Badelt

Joint work of Stefan Badelt, Ronny Lorenz, Ivo L. Hofacker

Main reference Stefan Badelt, Ronny Lorenz, Ivo L. Hofacker: “DrTransformer: Heuristic cotranscriptional RNA folding using the nearest neighbor energy model”, bioRxiv, Cold Spring Harbor Laboratory, 2022.

URL <https://doi.org/10.1101/2022.09.08.507181>

Cotranscriptional folding has the exciting potential to encode multiple functional important structures into a single molecule and visit them in a controlled manner. Unfortunately, the interpretation of experimental data on cotranscriptional folding is still heavily dependent on computational structure prediction, and it is easy to misinterpret data when using thermodynamic models. We use the stochastic simulator Kinfold, as well as a newly developed deterministic heuristic DrTransformer to show how cotranscriptional simulations can help with understanding experimental results and point out common mistakes in existing interpretations of data.

5.5 Eukaryotic riboswitch detection using inverse RNA folding

Danny Barash (Ben Gurion University - Beer Sheva, IL)

License © Creative Commons BY 4.0 International license
© Danny Barash

Joint work of Sumit Mukherjee, Matan Drory, Michelle Meyer, Danny Barash

The inverse RNA folding problem for designing sequences that fold into a given RNA secondary structure was introduced in the early 1990’s in Vienna. By an extension of this problem we use a coarse-grained approach to possibly detect novel eukaryotic riboswitches. The approach can tentatively be used for other domains and applications.

5.6 Design of RNA tandem repeats creating RNA droplets forming liquid-liquid phase separation

Sarah Berkemer (Ecole Polytechnique - Palaiseau, FR)

License © Creative Commons BY 4.0 International license
© Sarah Berkemer

Joint work of Sarah Berkemer, Ariel B Lindner, Yann Ponty, Carla Tous-Mayol

Various genetic disorders are caused by expansion of short tandem repeats as they aggregate in cells and form so-called RNA droplets or foci. However, the molecular mechanisms of RNA foci formation remains unclear. The aim of being able to design RNA tandem repeats and model RNA foci formation is twofold: it will help understand the mechanisms and therapies related to genetic disorders such as Huntington’s disease but at the same time serve as a method to spatial engineering inside cells as RNA droplets cause a liquid-liquid phase separation which can serve as process isolation and help to organize proteins and multienzyme pathways without fine-tuning RNA expression levels.

Phase-separating RNA molecule complexes are constructed from small repeating sequences, e.g. triplet repeats. Visualization of RNA foci is conducted by tagging droplets with e.g. GFP and corresponding adapters such as the MS2 aptamer.

Previous studies successfully showed the formation of RNA foci using various types of RNA triplet repeats and even longer repeat sequences where the formation of G-quadruplexes seems to be an important part for the interaction between two tandem repeat RNAs [1, 2, 3, 4].

Existing studies could experimentally show which RNA triplets are the most successful in forming RNA foci, however, the structure of RNA foci and their dynamics are not yet understood. Additionally, the liquid-liquid phase separation opens numerous possibilities for spatial engineering inside the cells, but we still lack the knowledge of structural and chemical properties of the RNA droplets and the space inside the foci. By designing RNA molecules that form droplets, we need to take into account interactions of more than two RNA sequences as well as possible interactions with binding proteins. Hence, we aim to develop design strategies for interacting short tandem RNA repeats and explore properties of RNA droplets and their formation.

References

- 1 Haotian Guo et al., <https://www.biorxiv.org/content/10.1101/2020.07.02.182527v2.full>
- 2 Jain and Vale, 2017, <https://doi.org/10.1038/nature22386>
- 3 Nguyen, Hori and Thirumalai 2021: <https://doi.org/10.1101/2021.02.20.432119>
- 4 Isiktas et al., 2022 <https://www.biorxiv.org/content/10.1101/2022.04.11.487960v1.full>

5.7 Scaling and Limits of DNA Strand Displacement Computing

Harold Fellermann (Newcastle University, GB)

License © Creative Commons BY 4.0 International license
© Harold Fellermann

Since it has been shown that DNA strand displacement (DSD) reactions can implement arbitrary chemical reaction networks, they have become a popular substrate for molecular computing applications. One question worth asking is whether and where there exists upper limits on the size and complexity of realistically achievable DSD circuits. To address this question, I am presenting as example application the design for a molecular QR code generator that displays a dedicated QR code for any configuration of n molecular input DNA strands. By increasing the number of inputs, the complexity of the circuit increases in a superexponential manner, as well as our current attempts to tame the number of required DSD gates that implement the required function. The second part of the talk presents experimental results on the scalability of DSD circuits and limits that arise from toehold occlusion or partially complementary toeholds. Motivation for this study is the realization that reversible toehold binding imposes an upper limit on toehold length of typically six to eight nucleotides. This in turn puts a hard limit on the number of distinct toeholds a circuit can employ. While the number of distinct nucleotide sequences is still quite large, crosstalk appears in system with significantly fewer toehold sequences once the sequence of supposedly distinct toeholds becomes similar enough to cause undesired interactions. We have systematically analyzed noise in DSD systems caused by crosstalk between signals with single and double mismatches. Our main result is that toehold occlusion might occur already in systems one order of magnitude below the theoretical upper limit to toehold domains.

5.8 Generative modelling of riboswitches with restricted Boltzmann machines

Jorge Fernández de Cossío Díaz (ENS - Paris, FR)

License  Creative Commons BY 4.0 International license

© Jorge Fernández de Cossío Díaz

Joint work of Jorge Fernández de Cossío Díaz, Andrea Di Gioacchino, Simona Cocco, Rémi Monasson, Bruno Sargueil, Yann Ponty, Bertrand Marchand, Pierre Hardouin, Francois-Xavier Lyonnet

Restricted Boltzmann machines (RBM) are energy-based latent variable generative models, consisting of two layers, that can offer interpretable representations of complex data. Recently they have been applied to modelling protein sequence data. In this talk, I will present evidence suggesting RBM are effective generative models of structured RNA. In particular, I consider the SAM riboswitch family, which regulates expression of downstream bacterial mRNAs by adopting competing structural conformations in response to the presence of a cellular metabolite. The RBM automatically infers relevant statistical features from the sequence data, such as conservation patterns, complementarity constraints consistent with the secondary structure, and the presence of a pseudoknot. The functionality of designed sequences has been validated experimentally by SHAPE mapping.

5.9 Get away from Plug and Pray: Synthetic Riboswitches - Applications and open Problems

Sven Findeiß (Universität Leipzig, DE)

License  Creative Commons BY 4.0 International license

© Sven Findeiß

Joint work of Sven Findeiß, Mario Mörl, Peter F. Stadler

I will talk about the collaborative projects with the group of Mario Mörl (Biochemistry department at Leipzig University) on transcription termination regulating riboswitches and how we put tRNA processing under ligand control. The presentation will summarize how the corresponding design models have been developed, implemented, and analyzed *in silico*, as well as the biochemical investigations *in vitro* and *in vivo*. I will not only show the success story but the main emphasis will be on the problems we faced, how we solved them, and especially the issues that remain.

References

- 1 Wachsmuth M. et al., 2013, <https://doi.org/10.1093/nar/gks1330>
- 2 Findeiß S. et al., 2018, <https://doi.org/10.1016/j.ymeth.2018.04.036>
- 3 Günzel C., et al., 2020, <https://doi.org/10.1080/15476286.2020.1816336>
- 4 Ender A. et al., 2021, <https://doi.org/10.1093/nar/gkaa1282>

5.10 Designing RNA during the DNA Origami revolution

Cody Geary (Aarhus University, DK)

License © Creative Commons BY 4.0 International license
© Cody Geary

Joint work of Cody Geary, Guido Grossi, Ewan K. S. McRae, Paul W. K. Rothemund, Ebbe S. Andersen
Main reference Cody Geary, Guido Grossi, Ewan K. S. McRae, Paul W. K. Rothemund, Ebbe S. Andersen: “RNA origami design tools enable cotranscriptional folding of kilobase-sized nanoscaffolds”. *Nat. Chem.* 13, 549–558 (2021).

URL <https://doi.org/10.1038/s41557-021-00679-1>

RNA is the punk-brother of DNA. While DNA plays by rules, RNA is more rebellious. The diverse structural features of RNA that make it a powerfully-functional molecule in biology also make it difficult to tame and rationally-design.

In contrast to engineered DNA nanostructures such as DNA origamis, natural RNA molecules in cells must fold under non-equilibrium conditions; the RNAs fold continuously while the strand is still emerging from the polymerase. While design of staple strands to produce DNA origami nanostructures can be easily automated by simple algorithms, producing a single-stranded RNA origami requires the entire sequence of the RNA to be designed by inverse folding, which is computationally much more challenging.

Our RNA design software ROAD begins with a random starting sequence, and over many iterations mutates that sequence to improve its folding into a target fold. ROAD uses both positive and negative design cycles to perform a gradient descent based on an adapting scoring function. The strategy is based on *in vitro* selection methods where the selection conditions gradually become more difficult over successive rounds.

5.11 Two design stories: probes for SARS-CoV-2 detection and CRISPR/Cas9 gRNAs

Jan Gorodkin (University of Copenhagen, DK)

License © Creative Commons BY 4.0 International license
© Jan Gorodkin

Joint work of Jan Gorodkin, Mohsen Mohammadniaei, Ming Zhang, Jon Ashley, Ulf Bech Christensen, Jan Friis-Hansen, Rasmus Gregersen, Jan Grom Lisby, Thomas Lars Benfield, Finn Erland Nielsen
Main reference Mohsen Mohammadniaei, Ming Zhang, Jon Ashley, Ulf Bech Christensen, Lennart Jan Friis-Hansen, Rasmus Gregersen, Jan Gorm Lisby, Thomas Lars Benfield, Finn Erland Nielsen, Jens Henning Rasmussen, Ellen Bøtker Pedersen, Anne Christine Rye Olinger, Lærke Tørring Kolding, Maryam Naseri, Tao Zheng, Wentao Wang, Jan Gorodkin, Yi Sun: “A non-enzymatic, isothermal strand displacement and amplification assay for rapid detection of SARS-CoV-2 RNA”. *Nat Commun* 12, 5089 (2021).

URL <https://doi.org/10.1038/s41467-021-25387-9>

Main reference Xi Xiang, Giulia I. Corsi, Christian Anthon, Kunli Qu, Xiaoguang Pan, Xue Liang, Peng Han, Zhanying Dong, Lijun Liu, Jiayan Zhong, Tao Ma, Jimbao Wang, Xiuqing Zhang, Hui Jiang, Fengping Xu, Xin Liu, Xun Xu, Jian Wang, Huanming Yang, Lars Bolund, George M. Church, Lin Lin, Jan Gorodkin, Yonglun Luo: “Enhancing CRISPR-Cas9 gRNA efficiency prediction by data integration and deep learning”. *Nat Commun* 12, 3238 (2021).

URL <https://doi.org/10.1038/s41467-021-23576-0>

I will present two design cases. The first case concern non-enzymatic isothermal strand displacement and amplification for rapid detection of SARS-CoV-2, which we accomplished through design of DNA probes that opens and binds to targeted locations of the SARS-CoV-2 genome. Through RNA folding considerations, we show why one of two probes are more successful and makes the detection possible. In the second case, design of CRISPR/Cas9 guide RNA (gRNA) are made from first generating cleavage efficiency data and subsequently train a deep learning-based neural network which has cutting-edge performance tested on independent data sets.

5.12 What can geometric combinatorics say about RNA design?

Christine Heitsch (Georgia Institute of Technology - Atlanta, US)

License © Creative Commons BY 4.0 International license
© Christine Heitsch

Joint work of Christine Heitsch, Svetlana Poznanović, et al.

Branching is a critical characteristic of RNA design, yet can be challenging to validate with thermodynamic optimization approaches. Using mathematical methods (convex polytopes and their normal fans), we can improve prediction accuracy on well-defined families while also illuminating why the general problem is so difficult.

5.13 Experiments in Deep Learning for RNA Secondary Structure Prediction

Ivo Hofacker (Universität Wien, AT)

License © Creative Commons BY 4.0 International license
© Ivo Hofacker

Joint work of Christoph Flamm, Julia Wielach, Michael T. Wolfinger, Stefan Badelt, Ronny Lorenz, Ivo L. Hofacker

Main reference Christoph Flamm, Julia Wielach, Michael T. Wolfinger, Stefan Badelt, Ronny Lorenz, Ivo L. Hofacker: “Caveats to Deep Learning Approaches to RNA Secondary Structure Prediction”, *Frontiers in Bioinformatics*, Vol. 2, 2022.

URL <https://doi.org/10.3389/fbinf.2022.835422>

Machine learning (ML) and in particular deep learning techniques have the potential to overcome shortcomings of current RNA secondary structure prediction methods, such as the inability to predict pseudoknots and poor treatment of non-canonical pairs. Several recent publications have proposed deep neural networks for RNA secondary structure prediction and reported excellent accuracies. However, these works build upon training sets that are derived from a relatively small number of RNA families and therefore do not properly represent the RNA structure space.

By folding random sequences using the RNAfold program of the ViennaRNA package, we can generate synthetic data sets that allow to test in detail which properties of the RNA folding map are easy or hard to learn for these networks. We find that structure features that are local in the base pairing matrix, such as stacks and interior loops, are easy to learn, while less local multi-loops are much harder. Most strikingly, the number of base pairs predicted by convolutional networks grows quadratically, rather than linearly, with sequence length.

Using inverse folding, we designed a further synthetic training set that contains the same structures as the widely used bprNA data set, and therefore exhibits the same lack of structure diversity in spite of near perfect randomness of the sequences. Networks trained on this data set achieve excellent performance on sequence that have no similarity to training sequences but fold into structures well represented in the training set. Nevertheless, the networks perform poorly on sequence folding into novel structures. This suggests, that the excellent performance reported in the literature is largely due biases in the data sets, i.e. that training and test sets that exhibit the same overrepresentation of a few well studied RNA families and their structures.

5.14 BarMap-QA – Cotranscriptional folding with quality assurance

Felix Kühnl (Universität Leipzig, DE)

License © Creative Commons BY 4.0 International license

© Felix Kühnl

URL https://www.bioinf.uni-leipzig.de/Software/BarMap_QA/

The structure of an RNA molecule is often a crucial characteristic to be able to explain its biological function. While studying the thermodynamically optimal (MFE) structure often yields important information, there are relevant cases where a computation of the MFE structure alone is not sufficient to understand a molecule's behaviour, for example in transcriptional riboswitches. Cotranscriptional folding simulations can thus be a helpful tool to gain a deeper understanding a given RNA.

In this talk, I present the software pipeline BarMap-QA, which relies on the BarMap framework by Hofacker et al., to simulate cotranscriptional folding of RNAs. The pipeline is not only very streamlined and easy to set up and run, but it also provides several quality measures to assess the quality of the conducted analysis and thus allow the user to optimally balance computational efficiency against simulation accuracy for a specific use case.

References

- 1 Kühnl F et al., accepted, <https://doi.org/10.1101/2020.01.06.895607>

5.15 Discovering RNA Self-Reproducers By In Silico And In Vitro Screening

Philippe Nghe (ESPCI - Paris, FR)

License © Creative Commons BY 4.0 International license

© Philippe Nghe


Joint work of Camille Lambert, Vaitea Opuu, Francesco Calvanese, Martin Weigt, Matteo Smerlak, Philippe Nghe

The RNA world hypothesis proposes that RNAs carry catalytic activity necessary for primordial evolution. A first necessary condition for evolution is reproduction. Whether self-reproduction is rare or common in the space of RNA sequences is central to assess the plausibility of this scenario. To date, two ribozymes have been shown to autocatalytically sustain their self-reproduction in the laboratory, starting from RNA oligomers: the Azoarcus ribozyme derived from the group I intron family (Hayden and Lehman 2006) and a fragmented ligase (Lincoln and Joyce 2009). In this project, we assess the probability of self-reproducing RNAs in sequence space by using as a starting point the Azoarcus ribozyme that can autocatalytically self-reproduce. We show that combining *in silico* and *in vitro* screening allows for the discovery of a large number of artificial self-reproducing ribozymes. For this, the strategy consists of: i) Identifying natural self-reproducing GIIIs; ii) Applying physics-based and machine learning methods to generate artificial candidates for self-reproduction; iii) Testing designed sequences for self-reproduction using high-throughput sequencing; v) characterizing the representative self-reproducers. We find that generative models that combine statistical signatures from pair correlations and secondary structure prediction are efficient at producing functional ribozymes more than 60 nucleotides away from the original sequence, whereas random mutations destroy activity after only a few.

These methods interpolate the natural diversity found in group I introns, from which self-reproducers can be successfully re-engineered. This overall shows that self-reproduction is not an exceptional property of a few laboratory-made RNAs, but is relatively widespread in the sequence space.

5.16 Physical modeling of RNA polymorphism

Samuela Pasquali (University Paris-Diderot, FR)

License  Creative Commons BY 4.0 International license
© Samuela Pasquali

Joint work of Samuela Pasquali, Konstantin Roeder

RNA molecules are characterized by the existence of a multitude of stable states that result in a frustrated energy landscape, where the observed structures depend sensibly on experimental conditions and can depend on the initial, unfolded, structure. Using both atomistic and coarse-grained physical models for RNAs, combined with enhanced sampling methods, we investigate the energy landscape of these systems to understand what are the most relevant structures in the different conditions. Using a few significant examples we show how the combination of these methods allowed us to rationalize the experimental evidence showing the concurrent existence of multiple states [1, 2]. The coarse-grained model we develop [3] is also a useful starting point to couple simulations with experimental data, moving toward integrative modeling. We have recently developed a simulation technique allowing to bias MD coarse-grained simulations with SAXS data on-the-fly [4], and a theoretical framework to perform fast constant pH simulations where we can model the system considering the exchange of charges with the solvent [5]. These developments allow us to account for the environment to obtain reasonable structures to then be studied more thoroughly with high-resolution modeling.

References

- 1 K Röder, G Stirnemann, AC Dock-Bregeon, DJ Wales, S Pasquali, Structural transitions in the RNA 7SK 5' hairpin and their effect on HEXIM binding, *Nucleic Acids Research* 48 (1), 373-389 (2020)
- 2 K Röder, AM Barker, A Whitehouse, S Pasquali, Investigating the structural changes due to adenosine methylation of the Kaposi's sarcoma-associated herpes virus ORF50 transcript, *PLOS Computational Biology* 2022, 18(5):e1010150, doi: 10.1371/journal.pcbi.1010150, PMID: 35617364
- 3 T. Cragolini, Y. Laurin, P. Derreumaux, S. Pasquali, The coarse-grained HiRE-RNA model for de novo calculations of RNA free energy surfaces, folding, pathways and complex structure predictions, *J. Chem. Theory Comput.*, 11, 3510 (2015)
- 4 L Mazzanti, L Alferkh, E Frezza, S Pasquali, Biasing RNA coarse-grained folding simulations with Small-Angle X-ray Scattering (SAXS) data, *BioXiv* doi: 10.1101/2021.03.29.437449 (2021)
- 5 S. Pasquali, E. Frezza, F.L. Barroso da Silva, Coarse-grained dynamic RNA titration simulations, *Interface Focus* 9: 20180066 (2019)

5.17 RNA dynamics: one basepair at a time

Katja Petzold (Karolinska Institute - Stockholm, SE)

License © Creative Commons BY 4.0 International license
© Katja Petzold

Joint work of Petzold, Katja and the entire PetzoldLab (alumni and current)

Many functions of RNA depend on rearrangements in secondary structure that are triggered by external factors, such as protein or small molecule binding. These transitions can feature on one hand localized structural changes in base-pairs or can be presented by a change in chemical identity of e.g. a nucleo-base tautomer [1]. We use and develop R1 ρ relaxation-dispersion NMR methods [2] for characterizing transient structures of RNA that exist in low abundance (populations <10%) and that are sampled on timescales spanning three orders of magnitude (μ s to s).

The characterization of transient structures in microRNA miR-34a targeting the mRNA of Sirt1 [3] will be discussed and a first glimpse into ribosomal dynamics will be provided. We have trapped these short-lived states and characterized their structure and impact on function.

References

- 1 I.J. Kimsey, K. Petzold, B. Sathyamoorthy, Z.W. Stein and H.M. Al-Hashimi Nature, 519 (7543), pp 315-320, 2015
- 2 J. Schlagnitweit, E. Steiner, H. Karlsson and K. Petzold, Chemistry – A European Journal; 24(23):6067-6070, 2018. M. Marušić, J. Schlagnitweit and K. Petzold, ChemBioChem 2019, 20, 2685-2710
- 3 L. Baronti, I. Guzzetti†, P. Ebrahimi†, S. Friebe Sandoz†, E. Steiner†, J. Schlagnitweit, B. Fromm, L. Silva, C. Fontana, A. Chen and K. Petzold, Nature 2020, 583, 139-144

5.18 Minimalistic RNA inverse folding

Yann Ponty (Ecole Polytechnique - Palaiseau, FR)

License © Creative Commons BY 4.0 International license
© Yann Ponty

Joint work of Yann Ponty, Jozef Hales, Alice Héliou, Jan Manuch, Ladislav Stacho, Sebastian Will, Stefan Hammer

Main reference Jozef Hales, Alice Héliou, Ján Manuch, Yann Ponty, Ladislav Stacho: “Combinatorial RNA Design: Designability and Structure-Approximating Algorithm in Watson-Crick and Nussinov-Jacobson Energy Models”, Algorithmica, Vol. 79(3), pp. 835–856, 2017.

URL <https://doi.org/10.1007/s00453-016-0196-x>

We consider two minimalistic instances of RNA Design, restricting our attention to a minimal instance of RNA inverse folding based on a simplified energy model, where any canonical base pair contributes equally to the free-energy. This greatly simplifies the study of algorithmic questions, hopefully enabling new (exact? efficient?) solutions to design problems that are usually approached in a heuristic fashion.


First, we consider the problem of counting/sampling sequences that are simultaneously compatible with a collection of secondary (2D) structures. Valid sequence assignments turn out to be in bijection, up to trivial symmetry, to independent sets of a compatibility graph, built as the union of base pairs from all structures. As all graphs can be obtained as unions of 2D structures, this implies #P-hardness of the counting problem. Yet, the problem can be solved using an DP algorithm that is fixed parameter tractable for the tree-width of

the graph. The associated algorithm can be further generalized to compute the partition function for generic constraints, and represents the engine of our declarative framework InfraRed for sequence sampling.

Next, we consider the inverse folding problem which starts from a single target 2D structure, and consists in finding a sequence that folds uniquely into the target with respect to base pair maximization. We first provide a complete characterization for designable structures without unpaired bases. More generally, we characterize extensive classes of (non-)designable structures, and prove the closure of the set of designable structures under the stutter operation. Finally, we consider a structure-approximating relaxation of the design, given a structure S (avoiding 2 basic undesignable motifs) transforms S into a designable structure by adding at most one base-pair to each helix. For all designable structures, a sequence can be generated in linear time, suggesting this relaxed version of design may be easier than the rigid version of the problem.

5.19 Challenges in designing RNA non-canonical modules

Vladimir Reinharz (University of Montreal, CA)

License  Creative Commons BY 4.0 International license
© Vladimir Reinharz

RNA structures depends largely on the geometry of interactions between its nucleotides. While the classical canonical/wobble interactions drive the folding of the major helices, there is a wide variety of different shapes that can connect through hydrogen bonds any nucleotide to any other. They have been classified by Leontis-Westhof into 12 non-canonical families. Graph algorithms have allowed to automatically retrieve all conserved network of non-canonical interactions in all known RNA structures. This work has exhibited the modularity and composability of these structures. Nonetheless, most of them don't have any associated thermodynamic parameter and it is still unknown if their folding is opportunistic or actually pushed for by these interactions. We ask as questions: What would be a rational scheme to design novel sequences folding in these shapes? How much of the context must be taken into account to ensure the correct folding? And how can chemical modifications enable unique modules?

5.20 Learning to Design RNA

Frederic Runge (Universität Freiburg, DE)

License  Creative Commons BY 4.0 International license
© Frederic Runge

Machine learning (ML) and especially deep learning (DL) approaches recently achieved remarkable results in different domains of life sciences. While such methods have entered many areas of molecular research, the field of RNA design still largely lacks deep learning-based approaches. To close this gap, we present two machine learning based approaches to tackle two different problems related to the field of RNA design. We present an automated deep reinforcement learning (AutoRL) approach that is capable of generating RNA sequences that fold into a desired secondary structure (inverse RNA folding) while often requiring only very few shots to yield a solution. Due to the sensitivity of deep RL algorithms to

their hyperparameter settings and the lack of similar work in the field, we use a meta-optimization approach to automatically find the best RL setting for solving the problem. Since inverse RNA Folding is fundamentally linked to RNA folding, we present a probabilistic Transformer for the secondary structure prediction problem. We show that our method outperforms previous work on a commonly used benchmark dataset from the literature and that it improves the quality of non-canonical base pair and pseudoknot predictions compared to previous work. Besides the advantages of a global reception due to self-attention compared to convolution neural networks, the probabilistic nature of our method allows to reconstruct structure ensembles learned from data.

5.21 Differential SHAPE probing to screen computationally designed RNA and to detect pseudoknot and non-canonical interactions

Bruno Sargueil (Paris Descartes University, FR)

License © Creative Commons BY 4.0 International license
© Bruno Sargueil

Joint work of Bruno Sargueil, Pierre Hardouin, Francois-Xavier Lyonnet, Elisa Frezza, Benoit Masquida, Yann Ponty, Sebastian Will, Simona Cocco, Rémy Monasson, Jorge Cossio, Andrea di Giocchino

Main reference Grégoire De Bisschop, Delphine Allouche, Elisa Frezza, Benoit Masquida, Yann Ponty, Sebastian Will, Bruno Sargueil: “Progress toward SHAPE Constrained Computational Prediction of Tertiary Interactions in RNA Structure”, *Non-Coding RNA*, Vol. 7(4), 2021.

URL <https://doi.org/10.3390/ncrna7040071>

The development of reliable RNA design processes requires experimental validation. RNA structure modelling from chemical probing experiments has made tremendous progress, however accurately predicting large RNA structures is still challenging for several reasons. In particular interactions such as pseudoknots and non-canonical base pairs which are not captured by the available incomplete thermodynamic model are hardly predicted efficiently. To identify nucleotides involved in pseudoknots and non-canonical interactions, we scrutinized the SHAPE reactivity of each nucleotide of a benchmark RNA under multiple conditions. We show that probing at increasing temperature was remarkably efficient at pointing to non-canonical interactions and pseudoknot pairings. The SHAPE probing technology was then use to screen for RNA computationally designed to interact with a small molecule

5.22 ENSnano: a 3D modeling software for designing complex DNA/RNA nanostructures

Nicolas Schabanel (ENS - Lyon, FR)

License © Creative Commons BY 4.0 International license
© Nicolas Schabanel

Joint work of Nicolas Levy, Allan Mills, Julie Finkel, Gaëtan Bellot, Nicolas Schabanel

URL <http://www.ens-lyon.fr/ensnano/>

Since the 1990s, increasingly complex nanostructures have been reliably obtained out of self-assembled DNA strands: from “simple” 2D shapes to 3D gears and articulated nano-objects, and even computing structures. The success of the assembly of these structures relies on a fine tuning of their structure to match the peculiar geometry of DNA helices. Various softwares have been developed to help the designer. These softwares provides essentially four kind of tools: an abstract representation of DNA helices (e.g. cadnano, scadnano,

DNAPen, 3DNA, Hex-tiles); a 3D view of the design (e.g., vHelix, Adenita, oxDNAviewer); fully automated design (e.g., BScOR, Daedalus, Perdix, Talos, Athena), generally dedicated to a specific kind of design, such as wireframe origamis; and coarse grain or thermodynamical physics simulations (e.g., oxDNA, MrDNA, SNUPI, Nupack, ViennaRNA,...). MagicDNA combines some of these approaches to ease the design of configurable DNA origamis.

We present our first step in the direction of conciliating all these different approaches and purposes into one single reliable GUI solution: the first fully usable version (design from scratch to export) of our general purpose 3D DNA nanostructure design software ENSnano. We believe that its intuitive, swift and yet powerful graphical interface, combining 2D and 3D editable views, allows fast and precise editing of DNA nanostructures. It also handles editing of large 2D/3D structures smoothly, and imports from the most common solutions. Our software extends the concept of grids introduced in cadnano; grids allows to abstract and articulated the different parts of a design. ENSnano also provides new design tools which speeds up considerably the design of complex large 3D structures, most notably: a 2D split view, which allows to edit intricate 3D structures which cannot easily be mapped in a 2D view, and a copy & repeat functionality, which takes advantage of the grids to design swiftly large repetitive chunks of a structure. ENSnano has been validated experimentally, as proven by the AFM images of a DNA origami entirely designed in ENSnano.

ENSnano is a light-weight ready-to-run independent single-file app, running seamlessly in most of the operating systems (Windows 10, MacOS 10.13+ and Linux), it thus does not require the installation of any other softwares such as Matlab, Maya or Samson. Precompiled versions for Windows and MacOS are ready to download on ENSnano website. In the coming months, we will add new features to our software to extend its capacities in the various directions discussed in this article. We decided to release now this first version of our software as its 3D and 2D editing interface is meeting our usability goals. Because of its stability and ease of use, we believe that ENSnano should find already its place in anyone's design chain, when precise editing of a larger nanostructure is needed.

Furthermore, we propose a new method for designing curved origamis that deviates radically from the pattern-based previous approaches. We have developed a new model for DNA double helices curved in 3D that allows us to directly position the DNA double helices constituting the desired shape in the 3D interface of our software ENSnano. The crossovers positions are then simply deduced from the 3D positions of the nucleotides, as predicted by our model. This geometry-based interactive approach shortcuts the tedious process of manually coming up with a pattern suited for the desired curvature, and furthermore allows to deal transparently with structure whose curvature varies continuously. We also propose an innovative 2D representation synchronizing curved parallel double helices without relying on insertions or deletions, by automatically adapting the cell width for each nucleotide in the array representation.

We provide experimental data validating our curvy DNA model by successfully annealing two DNA origamis conceived thanks to two new DNA design methods. The first origami consists in a 6-helices bundle following an interactively created bezier curve whose curvature gets as low as 4.7nm. The second is an asymmetrical Möbius torus whose DNA strands are routed along 2 spiraling helices covering its whole surface. This new spiraling technique, allowed by our DNA curvy model, enables to grasp crossovers within a continuous range which results in an easier-to-design and smoother surface. Both of our designs folded as is, without any need to redesign their xover schemes.

5.23 Single-Stranded Architectures for RNA Co-Transcriptional Folding

Shinnosuke Seki (The University of Electro-Communications - Tokyo, JP)

License © Creative Commons BY 4.0 International license
© Shinnosuke Seki

Joint work of Daria Pchelina, Nicolas Schabanel, Shinnosuke Seki, Guillaume Theyssier
Main reference Daria Pchelina, Nicolas Schabanel, Shinnosuke Seki, Guillaume Theyssier: “Oritatami Systems Assemble Shapes No Less Complex Than Tile Assembly Model (ATAM)”, in Proc. of the 39th International Symposium on Theoretical Aspects of Computer Science, STACS 2022, March 15-18, 2022, Marseille, France (Virtual Conference), LIPIcs, Vol. 219, pp. 51:1–51:23, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.
URL <https://doi.org/10.4230/LIPIcs.STACS.2022.51>

Oritatami (folding in Japanese) is a mathematical model of computation by co-transcriptional folding we proposed in 2016 and have been studying, primarily on its computational power. In this model, RNA co-transcriptional folding is generalized so that the bases (called “beads” herein) can be of arbitrarily defined, finitely-many types that may have arbitrary affinities with each other (rather than just the four bases in RNA with their fixed set of affinities), but restricted on the 2D plane. In this talk, we present the latest universal oritatami architecture that enables us to compute all computable functions (Turing universality) co-transcriptionally, with particular emphasis on simplicity of mechanisms it employs to read/write a bit, to store information, and to merge computational paths (erasure).

5.24 Coarse-grained modeling for RNA nanotechnology

Petr Sulc (Arizona State University - Tempe, US)

License © Creative Commons BY 4.0 International license
© Petr Sulc

Nucleic acid nanotechnology uses designed DNA or RNA strands that self-assemble into larger complexes and nanodevices. Computer modeling and simulations can provide crucial insights into function and design of such nanostructures. However, the sizes (up to thousands of base pairs) and timescales of their assembly (minutes to hours) of such nanodevices presents major challenge for modeling approaches. Here, we will present a coarse-grained model, oxDNA/oxRNA, specifically designed to simulate DNA and RNA nanotechnology, and we will demonstrate its application to RNA strand displacement reaction, a key mechanism in active nanotechnology devices which has recently been also identified to occur during RNA folding *in vivo*. We will then discuss applications of our modeling platform for inverse design of multicomponent nanostructure assemblies: how to design individual nucleic acid building blocks that self-assemble reliable into target multicomponent structure while avoiding kinetic traps and alternative free-energy minima? We show that through combination of multiscale modeling and mapping of the inverse design problem to Boolean Satisfiability Problem (SAT), it is possible to design nanostructures that assemble large-scale 3D assemblies, opening ways to use nucleic acids to biotemplated manufacturing.

5.25 Persuading tRNA to jump over stop codons

Andrew Torda (Universität Hamburg, DE)

License  Creative Commons BY 4.0 International license
© Andrew Torda

Joint work of Andrew E. Torda, Marco C. Matthies

Can one persuade a half-artificial tRNA to bind to a stop codon, pretend to be a tRNA-Ala and incorporate an alanine residue in a growing protein? If so, you might be on the way to alleviating a disease caused by an unwanted stop codon.

If you want to design an RNA sequence, you want a series of real nucleotides at the end of the day, but you may well go through some non-physical mixed states along the way. You can represent a base as some fraction of A plus C plus.. If you have an energy model, you can take the derivative of energy with respect to the composition at each site. This lets you use gradient-based methods to optimise your sequence.

We used the program DSS-Opt to find our artificial tRNA sequences, although this was no longer a de novo problem. The tRNA does not just have to fold correctly. It has to be able to convince an amino-acyl synthetase to charge it and then sneak past a host of recognition factors before a ribosome would consider taking it seriously. This means our calculations were far from de novo design. Only about 45% of the bases were actually optimised.

About half a dozen candidates were tested for charging by an alanine amino acyl-tRNA synthetase and then for stop-codon read-through with a luciferase assay. The winner of this was fed to an antibiotic-stalled ribosome and the structure solved by cryo-EM (acquisition code 7B5K). A bouncing baby half-designed tRNA smiled at the authors from the coordinates.

You could either view this as a triumph of design or you could say, less than half the sites in the molecule were actually chosen.

5.26 Kinetic features of RNA-RNA interactions

Maria Waldl (Universität Wien, AT)

License  Creative Commons BY 4.0 International license
© Maria Waldl

Interactions between RNAs are an essential mechanism in gene regulation. State-of-the-art computational genome-wide screens predict targets of regulatory RNAs based on thermodynamic stability but largely neglect kinetic effects. To overcome this limitation, we propose novel models of RNA-RNA interaction dynamics. On this basis we can improve our understanding of general principles that govern RNA-RNA interaction formation and improve target prediction tools.

While the dynamics of secondary structure formation of single RNAs have been successfully modeled using transition systems between conformations, analogous approaches for RNA-RNA interaction quickly lead to infeasibly large systems. Therefore, we propose reducing the interaction system to the direct trajectories (shortest paths) from possible first contacts to full hybridization. This key idea enables studying general principles and relevant features of the interaction formation as well as model details; e.g. the relative speed of intra- and intermolecular folding. Specifically, we isolate kinetic effects by comparing experimentally confirmed interactions from *Salmonella* and *E. coli* to a randomized background with similar thermodynamic properties.

These experiments indicate that native interactions are kinetically favored. Moreover, folding trajectories often look remarkably different depending on the site of the initial contact. Based on a machine learning classifier, we were able to identify a combination of interaction features that provide most information on the behavior of native RNA-RNA interactions. These features can be exploited to filter target predictions.

Due to the design of our RNA kinetics model, features like energy barriers can be computed efficiently. This enables refining genome-wide target predictions through kinetic criteria. Beyond these immediate practical improvements, we shed light on general principles like the long-debated influence of the accessibility of the initial contact site.

In the context of this seminar I would like to present this direct path models for interaction formation as well as the kinetic features that we identified and discuss how such features could extend current RNA design strategies.

5.27 Infrared: A sampling framework for RNA design... and beyond

Sebastian Will (Ecole Polytechnique - Palaiseau, FR)

License  Creative Commons BY 4.0 International license
© Sebastian Will

Joint work of Sebastian Will, Yann Ponty, Hua-Ting Yao

URL <http://www.lix.polytechnique.fr/~will/Software/Infrared/>

Infrared is a modeling framework for efficient targeted sampling and optimization. It was originally developed for implementing complex sequence design approaches with multiple objectives and side constraints, e.g. design of sequences with multiple RNA target structures while controlling the GC-content (RNARedPrint). Due to its declarative, compositional application programming/modeling interface, Infrared allows extending existing design tools to solve very specific design tasks, e.g. optimizing codon-usage while targeting RNA structures and (possibly) additional constraints. In the same way, it enables rapid development of completely new design tools like RNAPOND (and, due to its generality, even methods beyond design, e.g. alignment of RNAs with pseudoknots). A main feature of the system is its automatic adaptation to the complexity of the declaratively modeled task. For this purpose, the system implicitly derives fixed-parameter-tractable sampling and optimization algorithms using tree-decomposition. The talk outlines main properties and background of the system, its elementary usage, and presents concrete examples of design applications.

5.28 Forbidden RNA motifs and the cardinality of secondary structure space

Hua-Ting Yao (Universität Wien, AT)

License © Creative Commons BY 4.0 International license

© Hua-Ting Yao

Joint work of Hua-Ting Yao, Cédric Chauve, Mireille Regnier, Yann Ponty

Main reference Hua-Ting Yao, Cédric Chauve, Mireille Regnier, Yann Ponty: “Exponentially Few RNA Structures Are Designable”, in Proc. of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, BCB '19, p. 289–298, Association for Computing Machinery, 2019.

URL <https://doi.org/10.1145/3307339.3342163>

The problem of RNA design attempts to construct RNA sequences that perform a pre-defined biological function, identified by several additional constraints. One of the foremost objectives of RNA negative design is that the designed RNA sequence should adopt a pre-defined target secondary structure preferentially to any alternative structure, according to a given metrics and folding model. It was observed in several works that some secondary structures are undesignable, i.e. no RNA sequence can fold into the target structure while satisfying some criterion measuring how preferential this folding is compared to alternative conformations.

We show that the proportion of designable secondary structures decreases exponentially with the size of the target secondary structure, for various popular combinations of energy models and design objectives. This exponential decay is, at least in part, due to the existence of forbidden motifs, which can be generically constructed, and jointly analyzed to yield asymptotic upper bounds on the number of designable structures. Moreover, we define a lower bound of the structural ensemble defect. We show that, across uniformly distributed secondary structures, such a lower bound has a Normal limiting distribution with the expected value and the variance both linear to the size of the secondary structure.

Participants

- Ebbe Sloth Andersen
Aarhus University, DK
- Maciej Antczak
Poznan University of
Technology, PL
- Stefan Badelt
Universität Wien, AT
- Danny Barash
Ben Gurion University –
Beer Sheva, IL
- Sarah Berkemer
Ecole Polytechnique –
Palaiseau, FR
- Anne Condon
University of British Columbia –
Vancouver, CA
- Harold Fellermann
Newcastle University, GB
- Jorge Fernández de Cossío
Díaz
ENS – Paris, FR
- Sven Findeiß
Universität Leipzig, DE
- Christoph Flamm
Universität Wien, AT
- Cody Geary
Aarhus University, DK
- Jan Gorodkin
University of Copenhagen, DK
- Christine Heitsch
Georgia Institute of Technology
– Atlanta, US
- Ivo Hofacker
Universität Wien, AT
- Felix Kühnl
Universität Leipzig, DE
- István Miklós
ELKH – Budapest, HU
- Philippe Nghe
ESPCI – Paris, FR
- Cyrille Merleau Nono Saha
MPI für Mathematik in den
Naturwissen. – Leipzig, DE
- Samuela Pasquali
University Paris-Diderot, FR
- Katja Petzold
Karolinska Institute –
Stockholm, SE
- Yann Ponty
Ecole Polytechnique –
Palaiseau, FR
- Vladimir Reinharz
University of Montreal, CA
- Lorenz Ronny
Universität Wien, AT
- Frederic Runge
Universität Freiburg, DE
- Bruno Sargueil
Paris Descartes University, FR
- Nicolas Schabanel
ENS – Lyon, FR
- Shinnosuke Seki
The University of
Electro-Communications –
Tokyo, JP
- Petr Sulc
Arizona State University –
Tempe, US
- Marta Szachniuk
Poznan University of
Technology, PL
- Andrew Torda
Universität Hamburg, DE
- Maria Waldl
Universität Wien, AT
- Sebastian Will
Ecole Polytechnique –
Palaiseau, FR
- Hua-Ting Yao
Universität Wien, AT

