Quantitative modeling of single-molecule RNA force-extension experiments

Ralf Bundschuh
Department of Physics, Ohio State University

Collaborators:
Ulrich Gerland, LMU München
Terence Hwa, UC San Diego

Outline:
• Introduction to RNA
• Single-molecule experiments
• Quantitative description of force-induced denaturation
• Why there is no structure in force-extension curves
• How we can still see some structure
• Conclusions and outlook
RNA is a heteropolymer of four different bases G, C, A, and U.

Primary structure: Sequence, e.g.,
GCGGAAUUGCAGUUGGAGAGCGCCAGACUGAAAUUCUGGAGGGUCUGUGUUCGAUCCACAGAAUUCGCACCA

Strongest interaction: Watson-Crick base pairing (G–C and A–U)
→ secondary structure

Spatial arrangement → tertiary structure

Structure encoded in sequence

Structure determines biological function
Single-molecule experiments

- RNA secondary structure difficult to determine experimentally since RNA is very floppy \( \Rightarrow \) no two copies of an RNA have the same shape

- Alternative: single-molecule experiments
  - Attach ends of RNA molecule to two beads
  - Keep beads at fixed distance \( R \) with optical tweezers
  - Measure force \( f \) on beads as function of distance \( R \)
  - P5ab hairpin of tetrahymena thermophila group I intron

Liphardt, Onoa, Smith, Tinoco, and Bustamante, Science, 2001

- Method to learn something about secondary structure?
Quantitative prediction I

- Experiments hard ⇒ quantitative modeling
- Two ingredients: secondary structure and polymer physics of backbone
- Backbone physics
  - Elastic freely jointed chain
  - Persistence length 1.9nm/base distance 0.7nm
  - Partition function at distance $R$ and length $m$: $W(R; m)$
- Secondary structure:
  - Need base pairing partition function $Q(m)$ of RNA molecule given
    $m$ “exterior” bases
- Total partition function:
  \[
  Z(R) = \sum_{m=0}^{N} Q(m) W(R; m)
  \]
  ⇒ Complete knowledge of the thermodynamics
How to get secondary structure partition function $Q(m)$ of RNA molecule given $m$ “exterior” bases?

Consider all possible sets of base pairs formed

Neglect pseudo-knots

Diagramatic representation

Assign energy $E[S] = \sum_{(i,j) \in S} \varepsilon_{i,j}$ to each structure $S$

Total partition function

$$Z = \sum_{m} Q(m) = \sum_{\{S\}} \exp(-E[S]/T)$$

Partition function generated exactly by Hartree equation

$O(N^3)$ algorithm for exact partition function (McCaskill, Biopolymers 29, 1990)
• As byproduct calculates partition function $Z_{i,j}$ for substrand from base $i$ to base $j$

• Introduce partition function $Q_j(m)$ for first $j$ bases with $m$ exterior bases

• Fulfills recursion equation

$$Q_j(m) = Q_{j-1}(m - 1) + \sum_{k=1}^{j-1} Q_{k-1}(m) e^{-\beta \epsilon_{k,j}} Z_{k+1,j-1}$$

• Can be calculated by modifying “Vienna package”


• Uses very detailed free energy rules

Quantitative prediction IV

- Apply to P5ab hairpin of tetrahymena thermophila group I intron

- **Quantitative prediction**

- **Disadvantages compared to experiment:**
  - No tertiary structure or pseudo-knots
  - Parameters not exactly known

- **Advantages over experiment:**
  - Can be rapidly applied to arbitrary sequence
  - Intermediate structures can be investigated

- Offered as interactive web server
  
  [http://bioserv.mps.ohio-state.edu/rna](http://bioserv.mps.ohio-state.edu/rna)
**FEC structure I**

- Apply to full *tetrahymena thermophila* group I intron

- Group I intron contains **pseudo-knot**!

- Quantitative modeling ignores pseudo-knot
  \[\Rightarrow\] known **inactive** conformation


- **No sign** of secondary structure!

- Consistent with experiments on single-stranded DNA

  Maier, Strick, Croquette, and Bensimon, *Single Molecules*, 2000
FEC structure II

• What’s happening?

• Look at intermediate structure (here $R = 100nm$)

• “Socks on the clothes line”

• Compensation effect:
  – Extension $R$ is increased
  – One of the “socks” disappears
  – The other “socks” take up the slag

⇒ No rapid change in force as sock disappears
⇒ smooth force-extension curve
• What can be done?

• Add spring

• Spring constant $\lambda$

• Vary $R_t$

• Measure extension $\langle R \rangle$ and fluctuations

\[
\delta R \equiv \sqrt{\langle (R - \langle R \rangle)^2 \rangle}
\]
• Fluctuations show structure

• Peaks in fluctuations correlated with disappearance of secondary structure elements
Structure by pulling III

- Alternative experimental setup: pulling through a nano-pore

![Diagram](image)

- Kinetic model: diffusion in a time dependent one-dimensional energy landscape ⇒ Monte Carlo simulation

![Graph](image)

- Signature of every structural element
Structure by pulling IV

- Extract sequence position of stalling sites
- Repeat experiment in reverse direction

- Match stalling sites by sequence comparison
  ⇒ Reconstruct structure
Structure by pulling V

- Overall structure reconstructed correctly

- Can distinguish different structures on the same sequence

- Even pseudoknot reconstructed

- “Just” needs to be implemented experimentally . . .
**Conclusions and outlook**

**Conclusions:**

- **Quantitative description of single-molecule experiments possible**
- **Force-extension curves do not reveal secondary structure information due to compensation effects**
- **Single-molecule experiments reveal structure information through measurements of fluctuations or with the help of a nano-pore**

**Challenges:**

- **Comparison with experiments on medium-sized RNA**
- **Include protein–RNA interactions**
- **Kinetics**