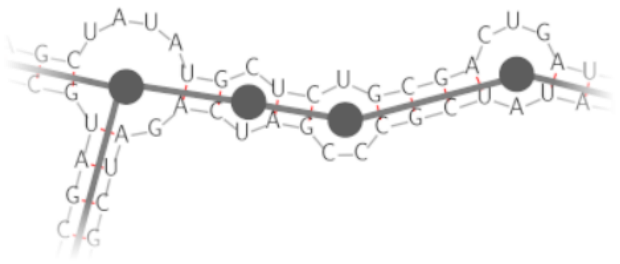
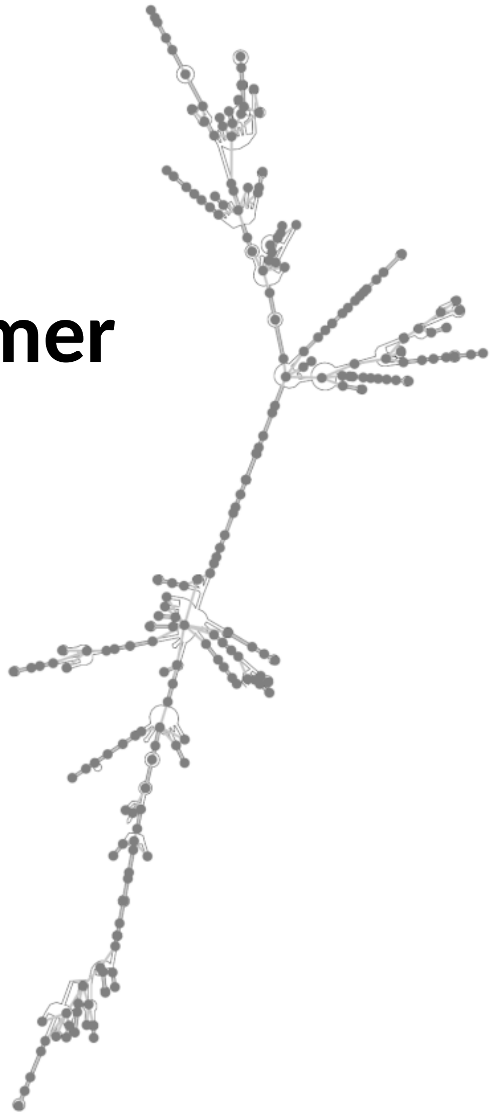


RNA as a randomly branched polymer



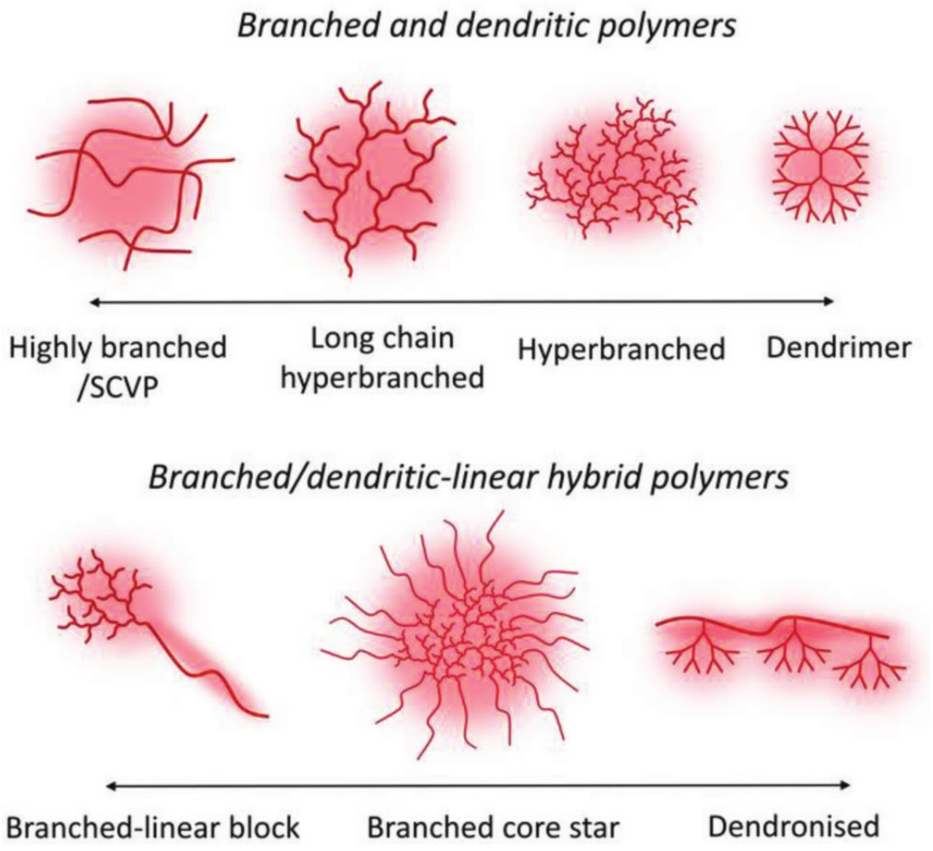
Anže Božič

Department of Theoretical Physics, IJS, Ljubljana

anze.bozic@ijs.si

Structural and topological properties of RNA in living systems,
Trento, 30. 01.—02. 02. 2023

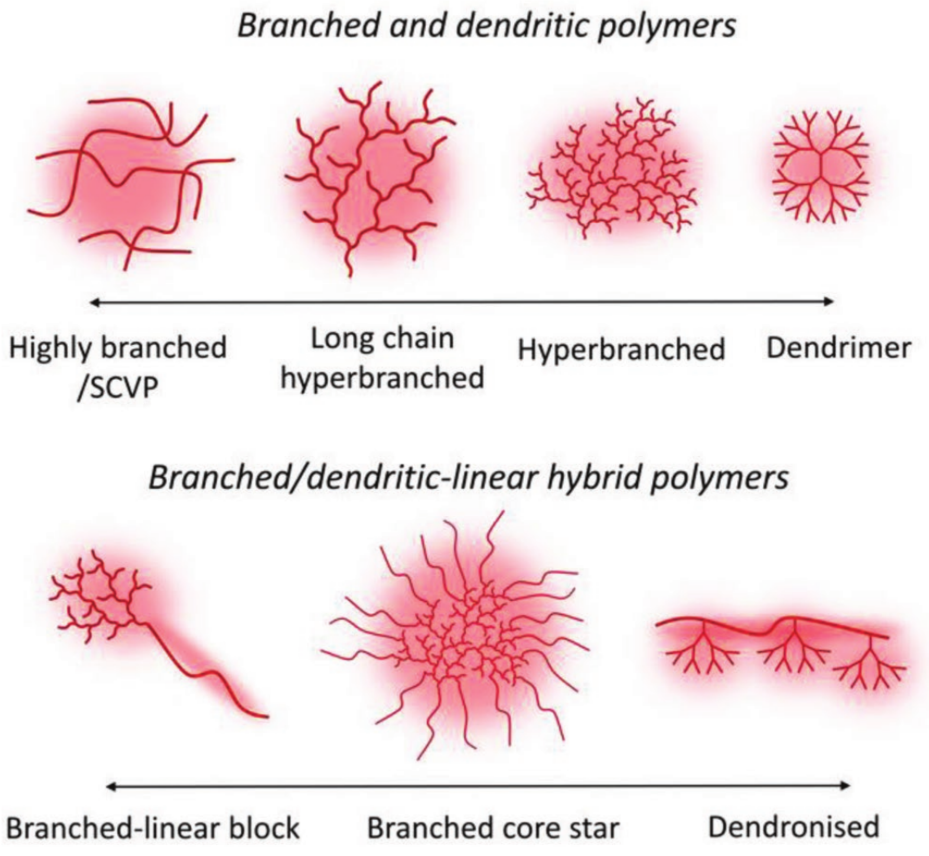
Branched polymers



AB Cook & S Perrier, Adv Func Mat 2020

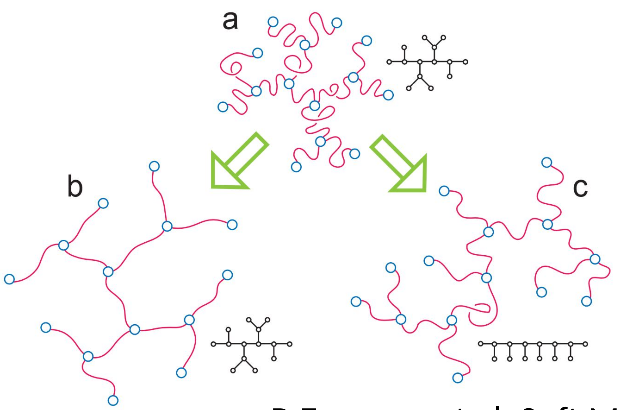
- polymers, ceramic aggregates, polymeric networks, gels, biopolymers
- favourable properties: high surface functionality, globular conformation, high solubilities, …
- two broad classes: regular and randomly branched/branching

Branched polymers



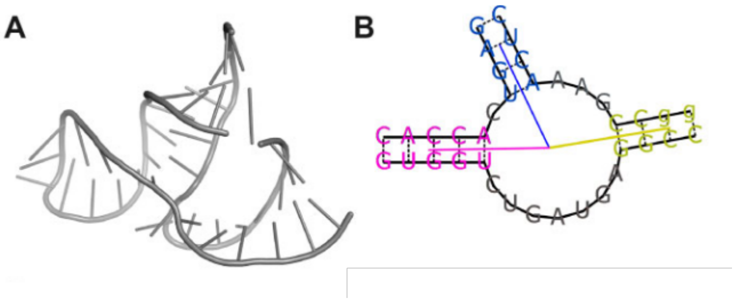
AB Cook & S Perrier, Adv Func Mat 2020

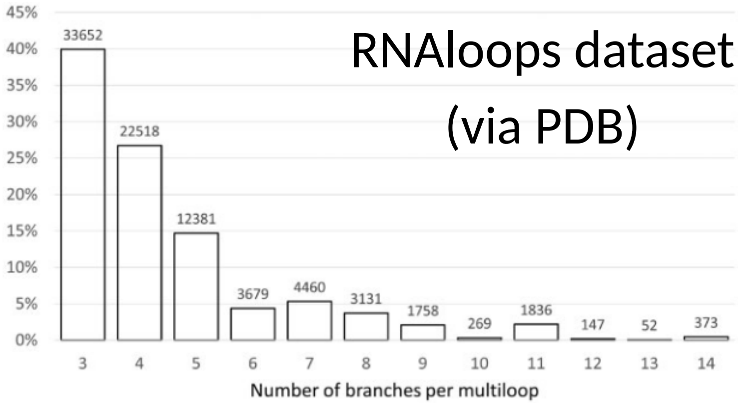
- polymers, ceramic aggregates, polymeric networks, gels, biopolymers
- favourable properties: high surface functionality, globular conformation, high solubilities, ...
- two broad classes: regular and randomly branched/branching



R Everaers et al, Soft Matter 2017

Branching structure of RNA

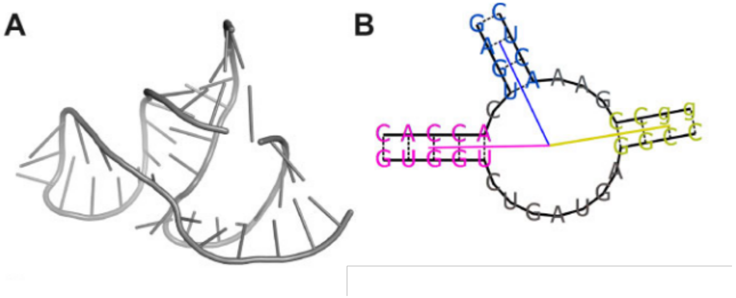


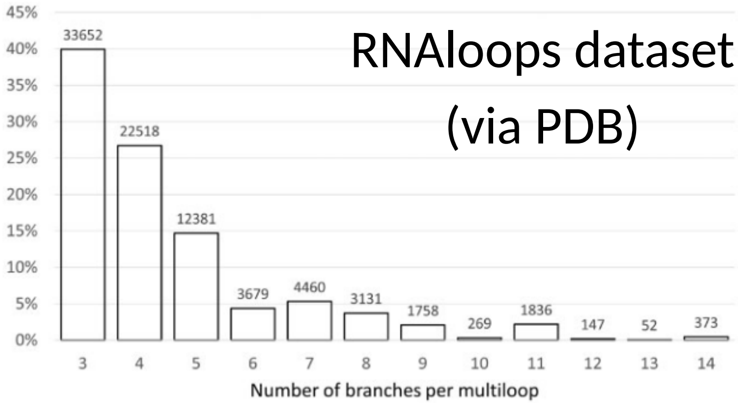


RNAloops dataset
(via PDB)

J Wiedemann et al, Bioinf 2022

Branching structure of RNA



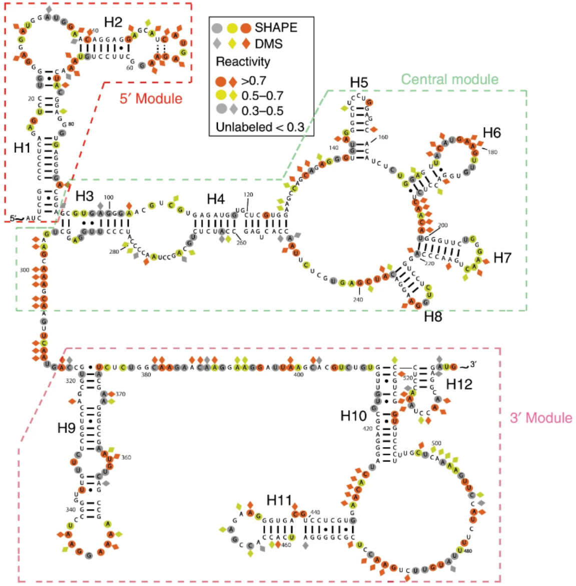


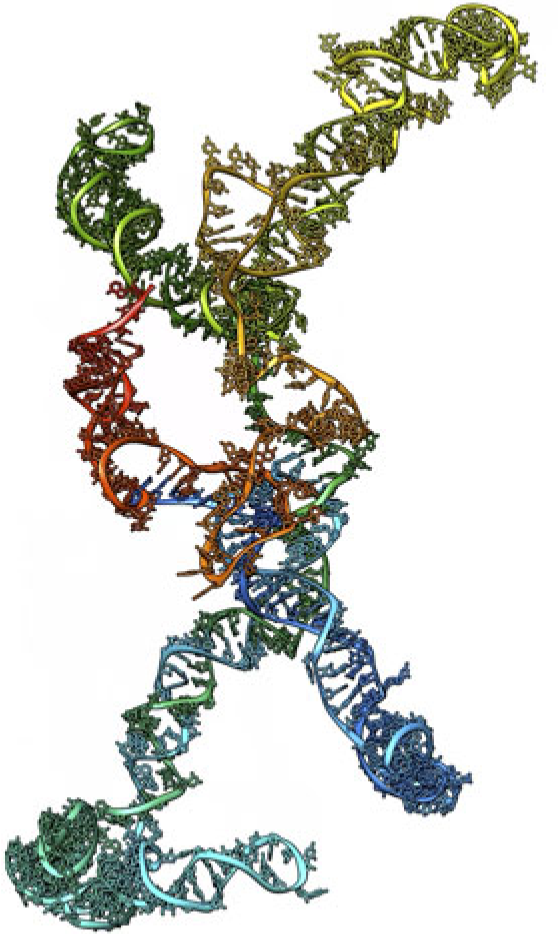
RNAloops dataset

(via PDB)

J Wiedemann et al, Bioinf 2022

"Braveheart" lncRNA (~ 600 nt)

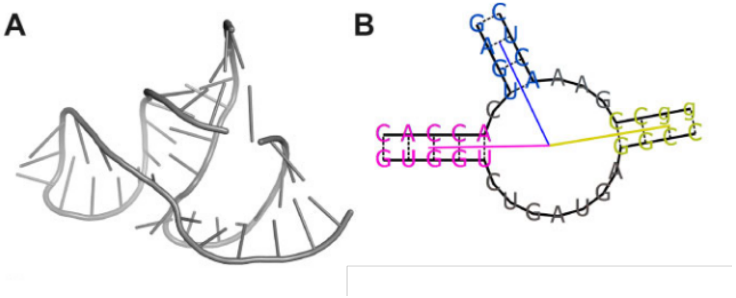


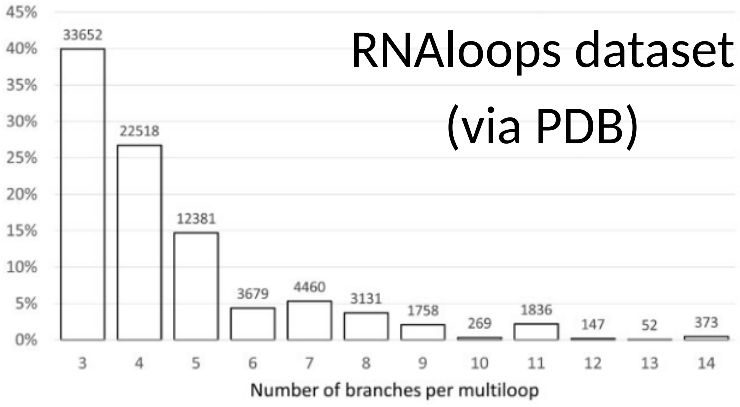


DN Kim et al, Nat Comm 2020

Branching structure of RNA

“Braveheart” lncRNA (~ 600 nt)

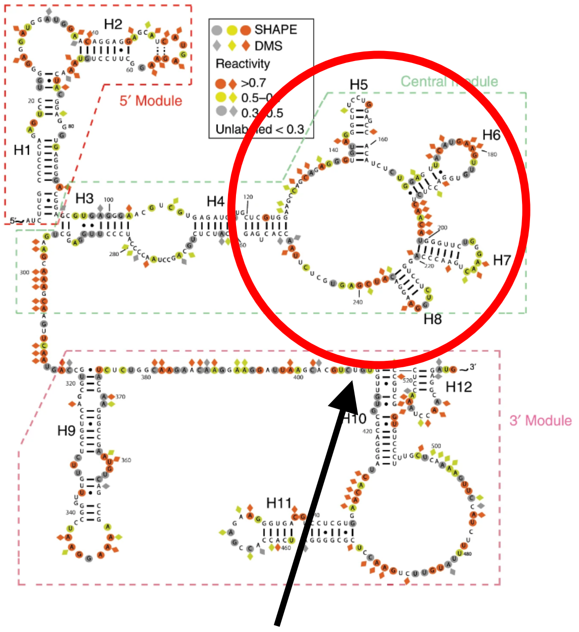




RNAloops dataset

(via PDB)

J Wiedemann et al, Bioinf 2022

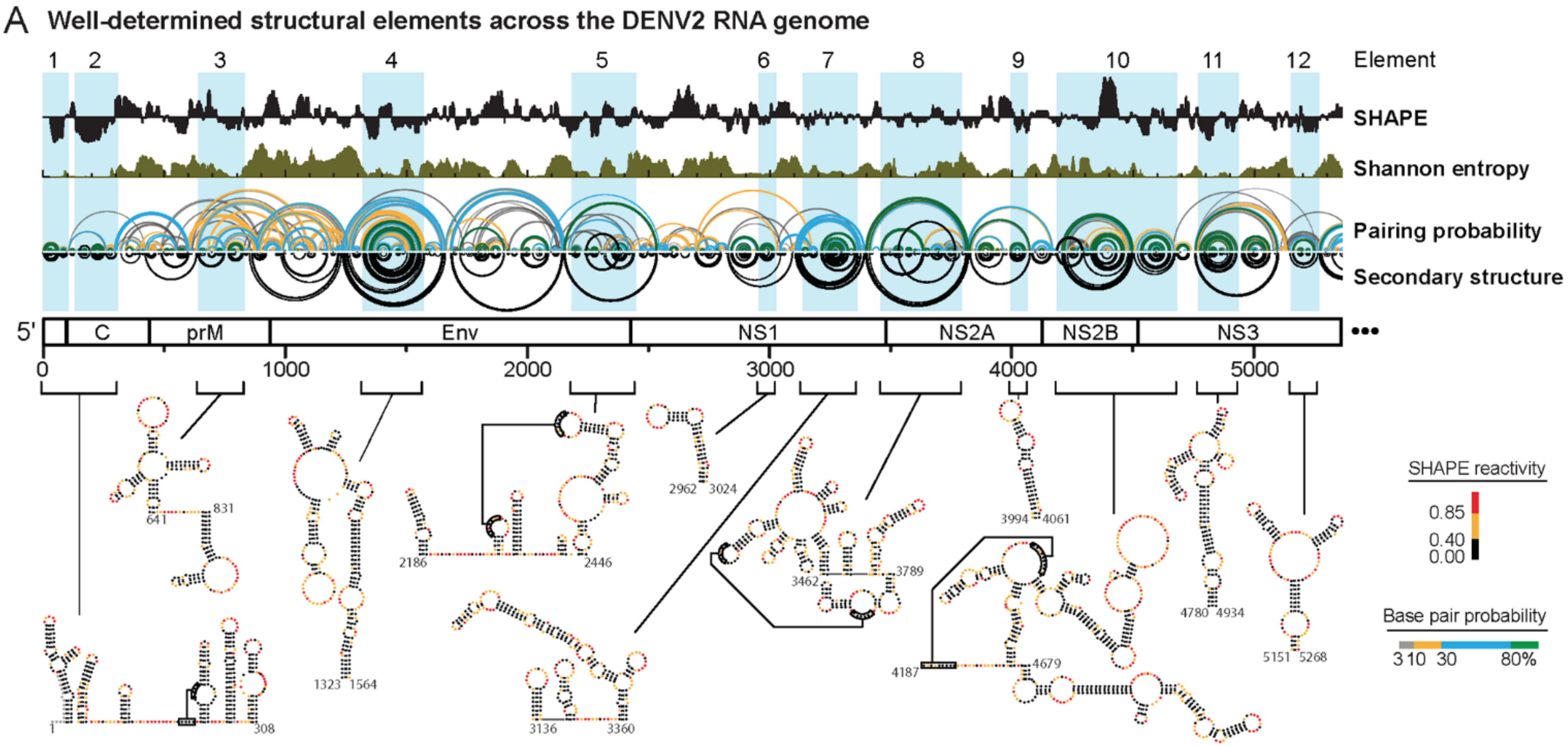


multiloop
of degree 5

DN Kim et al, Nat Comm 2020

Branching structure of RNA

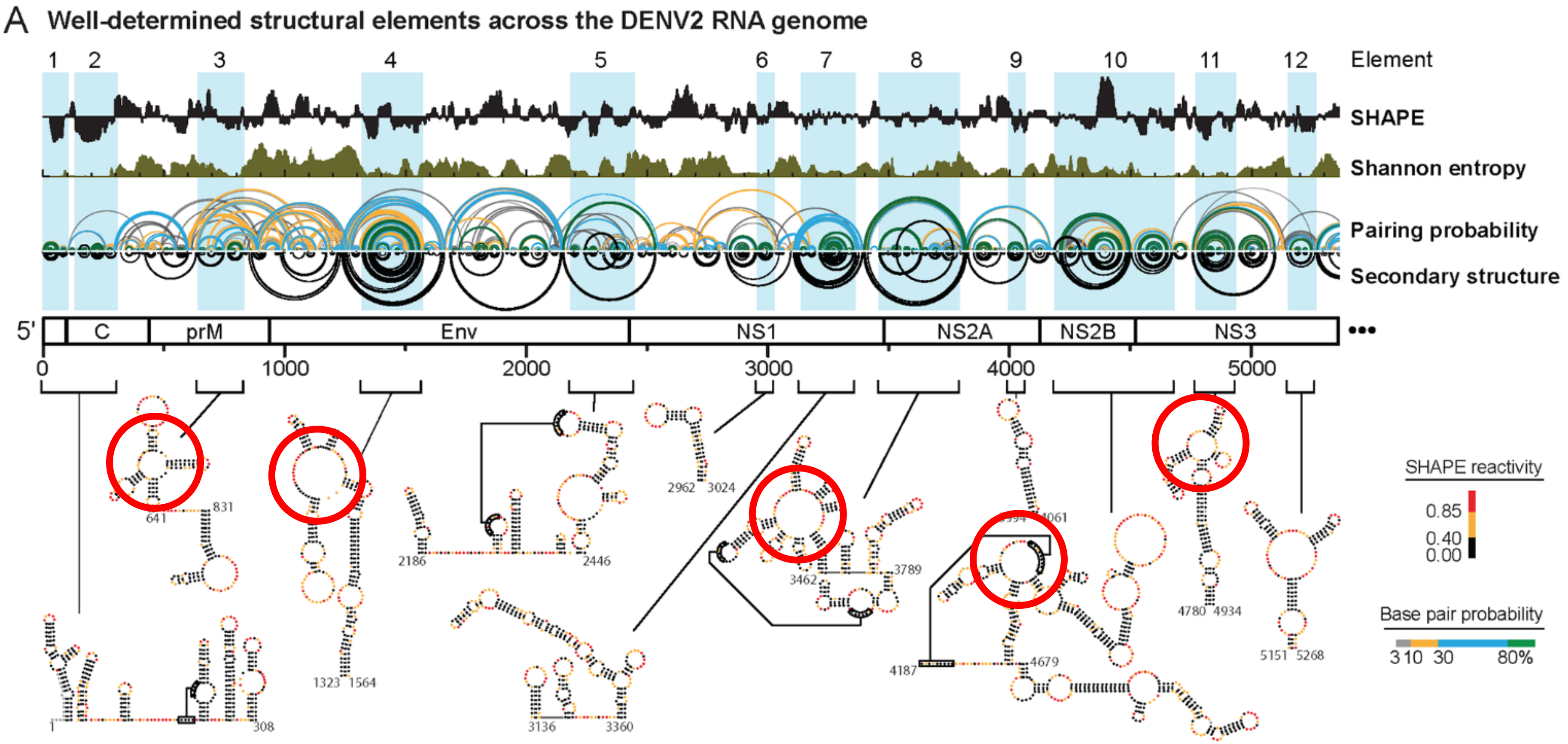
Dengue virus gRNA (~ 10700 nt)



KN Weeks, Acc Chem Res 2021

Branching structure of RNA

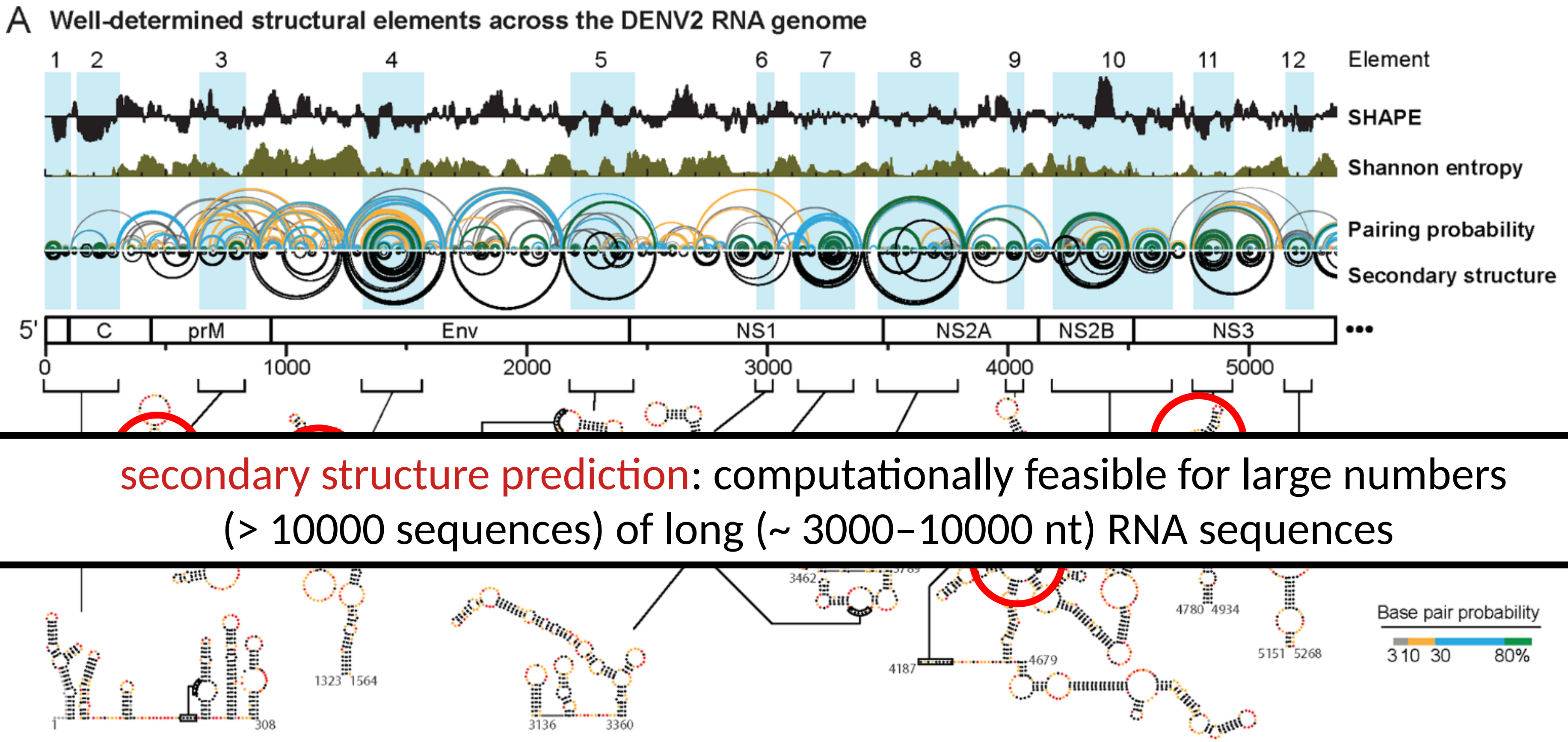
Dengue virus gRNA (~ 10700 nt)



KN Weeks, Acc Chem Res 2021

Branching structure of RNA

Dengue virus gRNA (~ 10700 nt)

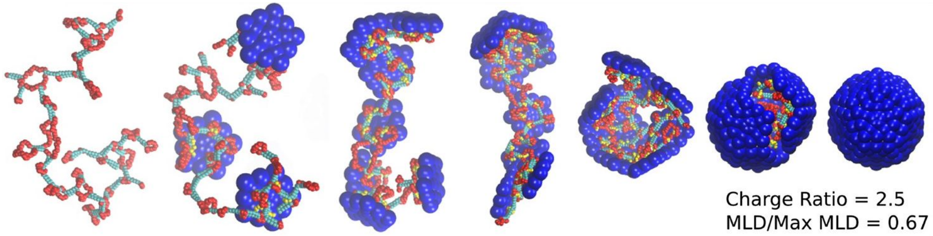


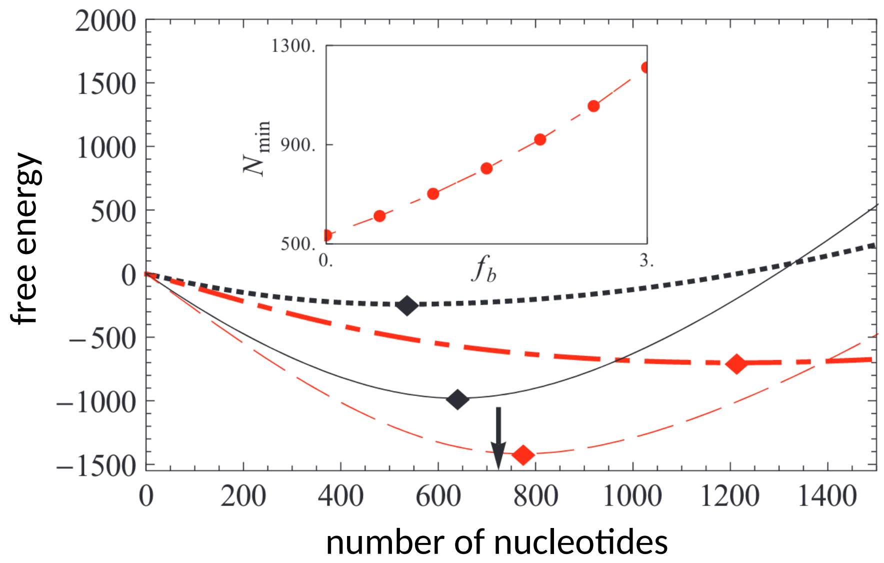
secondary structure prediction: computationally feasible for large numbers (> 10000 sequences) of long (~ 3000–10000 nt) RNA sequences

KN Weeks, Acc Chem Res 2021

Branching structure of *viral* RNA

JD Perlmutter et al, elife 2013

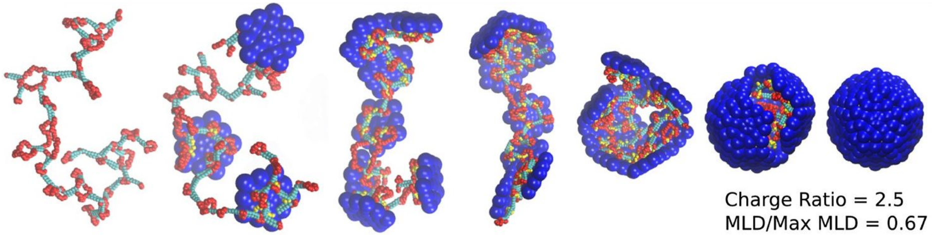


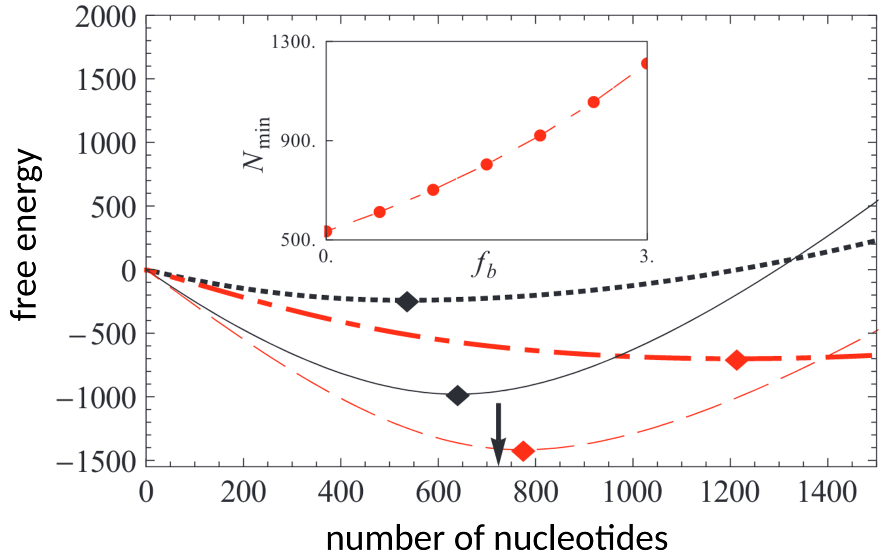


G Erdemci-Tardogan et al, Phys Rev E 2014

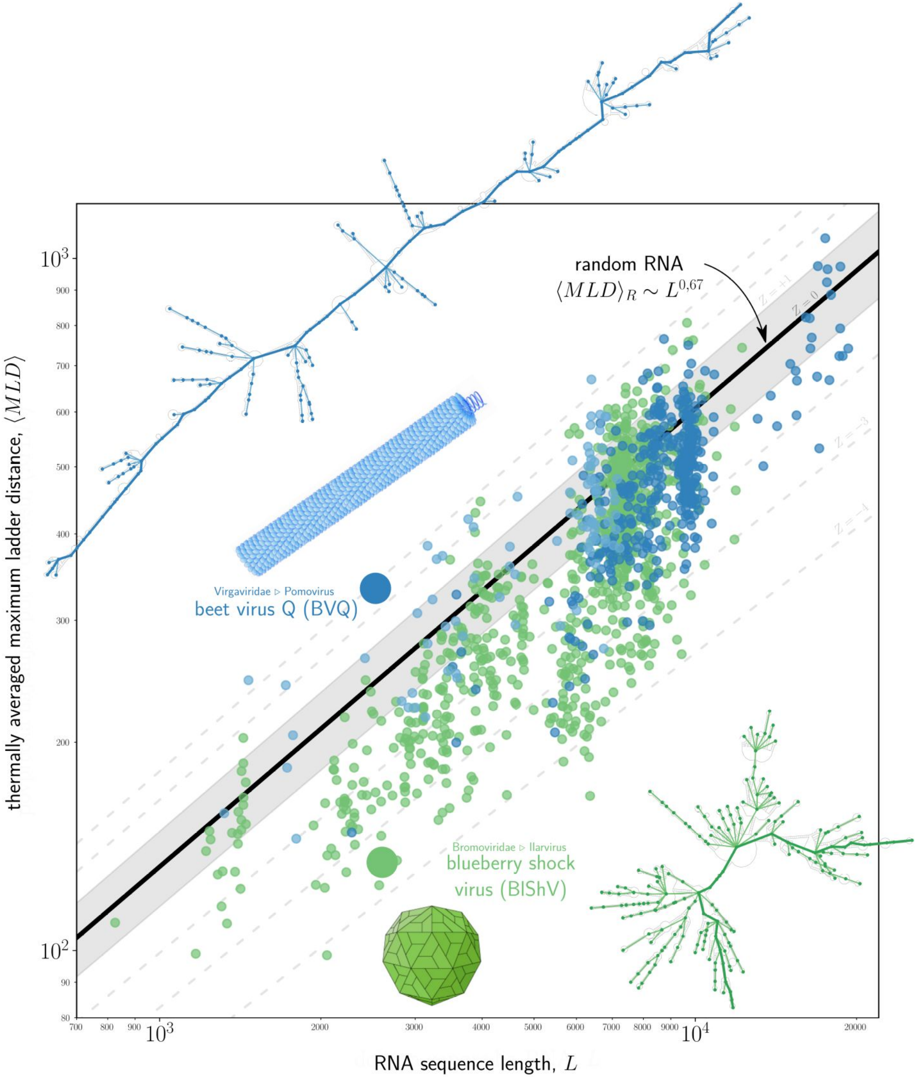
Branching structure of *viral* RNA

JD Perlmutter et al, elife 2013





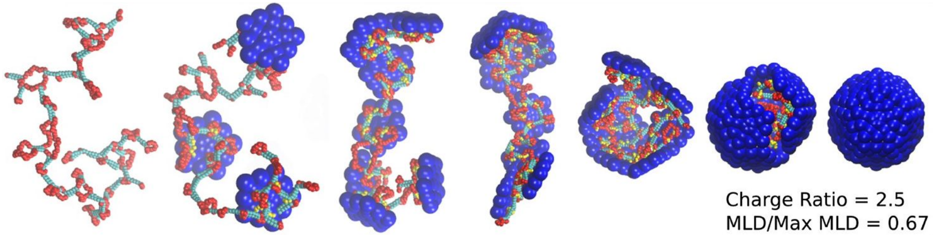
G Erdemci-Tardogan et al, Phys Rev E 2014

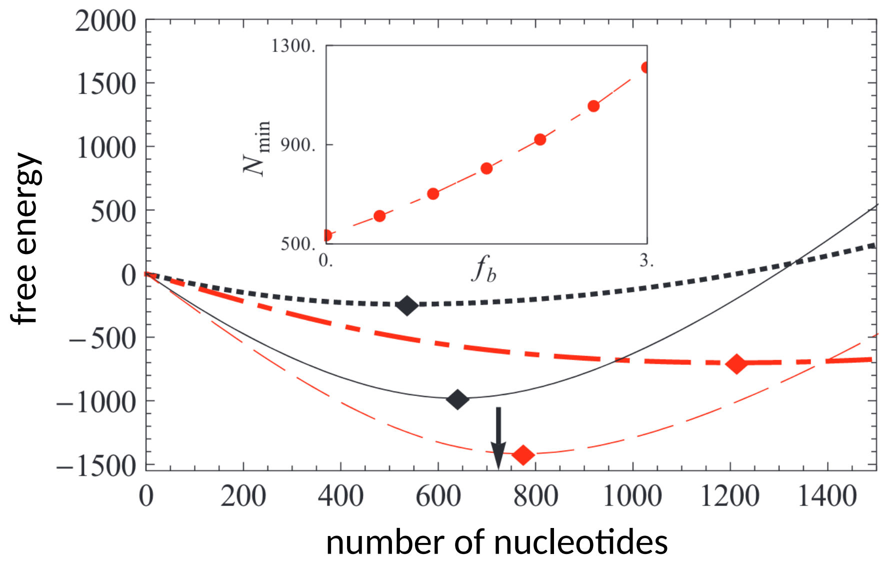


Vaupotič et al, arXiv:2212.00829

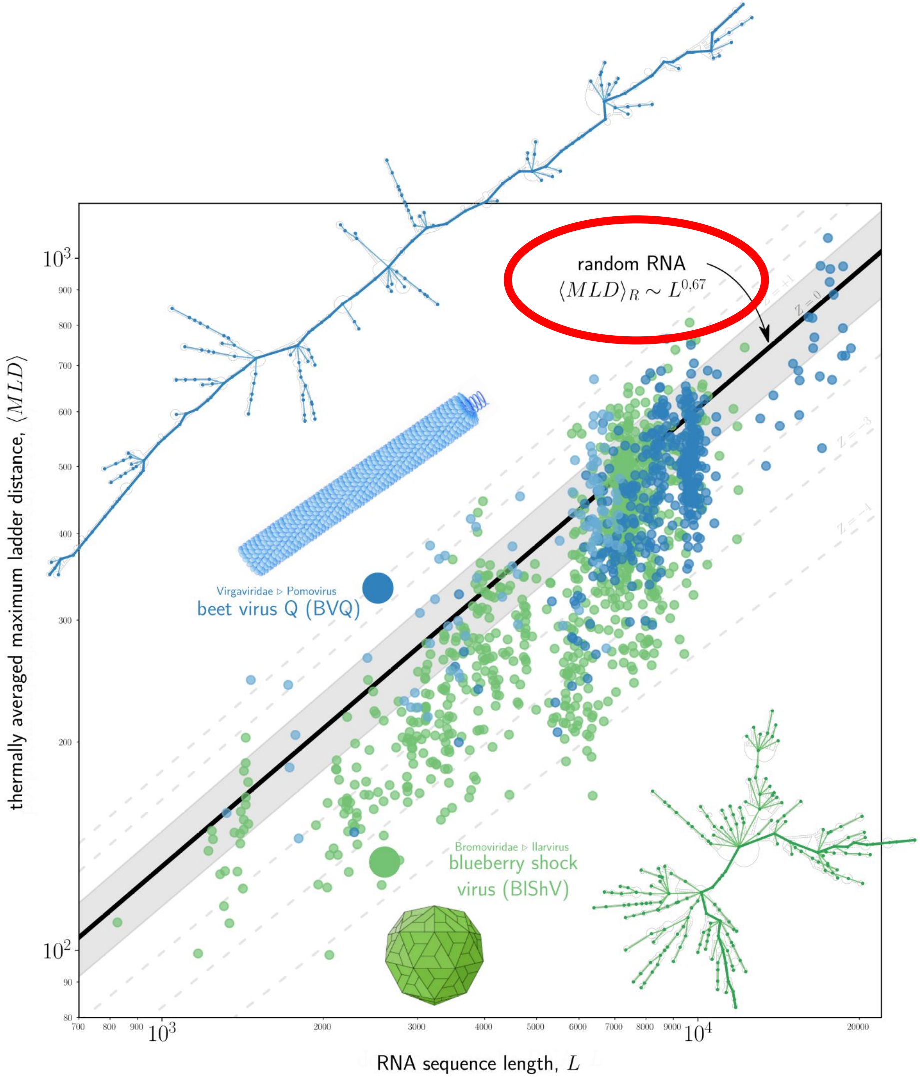
Branching structure of *viral* RNA

JD Perlmutter et al, elife 2013





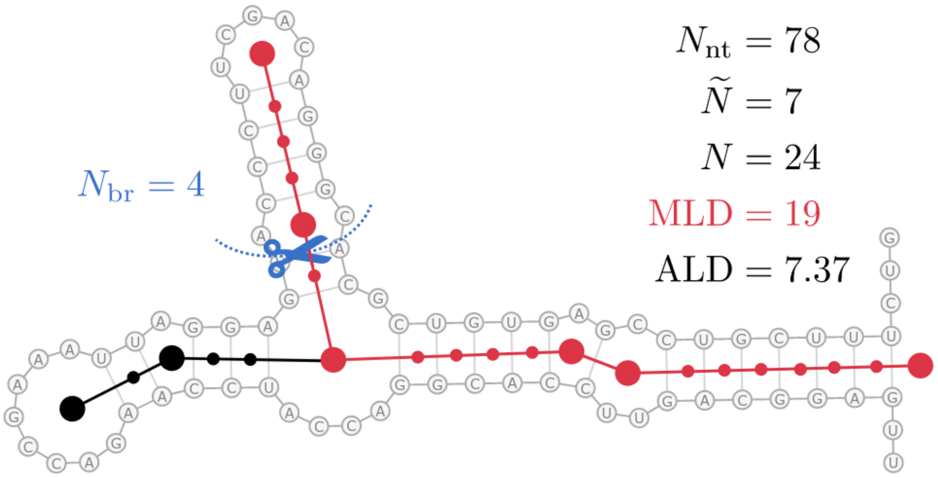
G Erdemci-Tardogan et al, Phys Rev E 2014

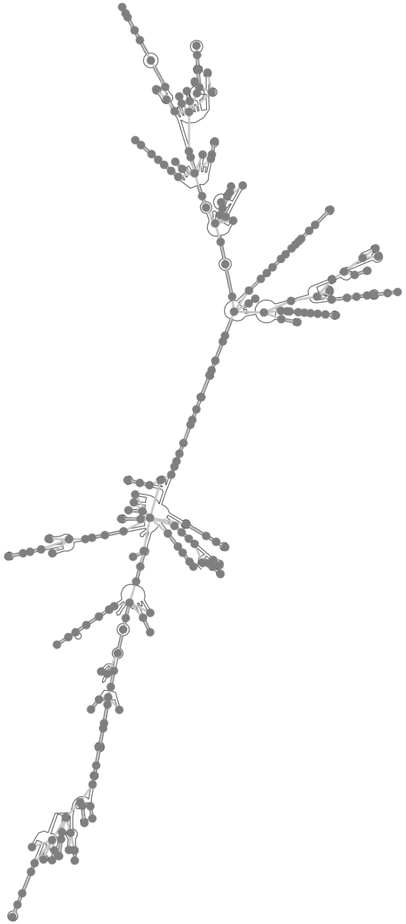


Vaupotič et al, arXiv:2212.00829

RNA as a branched polymer

RNA-as-graph: paired (double-stranded) regions as weighted edges and single-stranded regions as nodes (N)



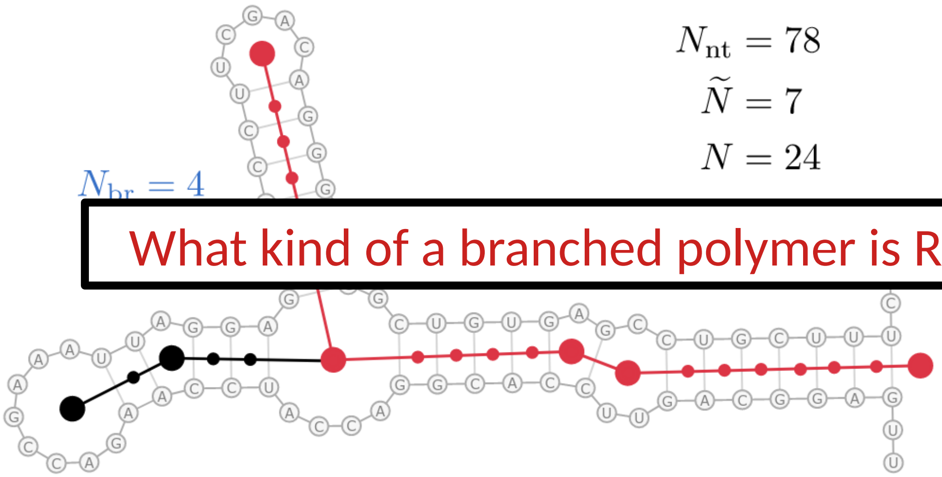


analysis of topological and physical properties
(node degree distribution, path distribution, …)

See also: T Schlick, Methods 2018

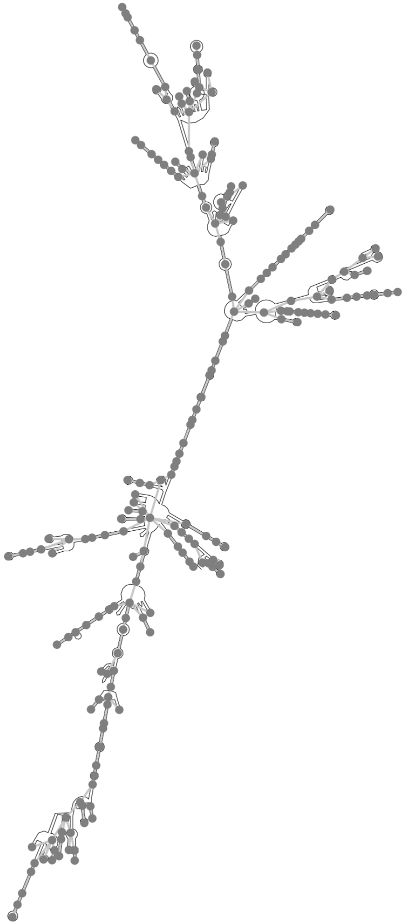
RNA as a branched polymer

RNA-as-graph: paired (double-stranded) regions as weighted edges and single-stranded regions as nodes (N)



$$N_{\rm nt} = 78$$
$$\widetilde{N} = 7$$
$$N = 24$$

What kind of a branched polymer is RNA?



analysis of topological and physical properties (node degree distribution, path distribution, …)

See also: T Schlick, Methods 2018

Scaling of branched polymers

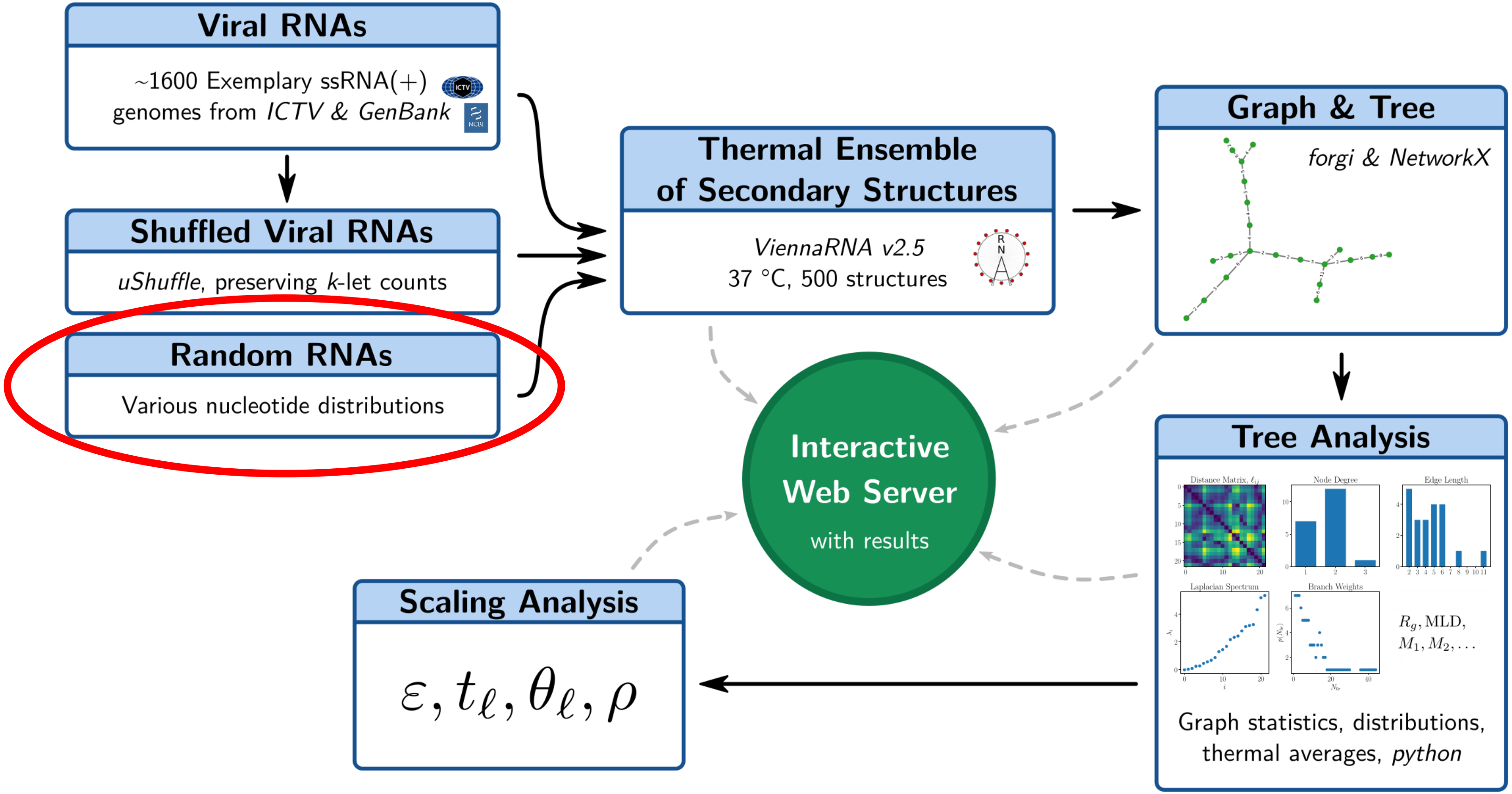
- description at the level of statistical mechanics (thermal ensembles of RNA structures)
- physical (linear) size not enough for branched polymers → topology of branching
- two topological measures: average length of linear paths on the tree (**exponent ρ**) and average branch weight (**exponent ε**)
- Flory theory: **topological measures as a proxy of polymer size**

$$\begin{aligned} \langle R_g^2 \rangle &\sim N^{2\nu} \quad \text{(polymer size)} \\ \langle \mathrm{MLD} \rangle &\sim N^{\rho} \quad \text{(polymer extension)} \\ \langle N_{br} \rangle &\sim N^{\varepsilon} \quad \text{(branch size)} \end{aligned}$$

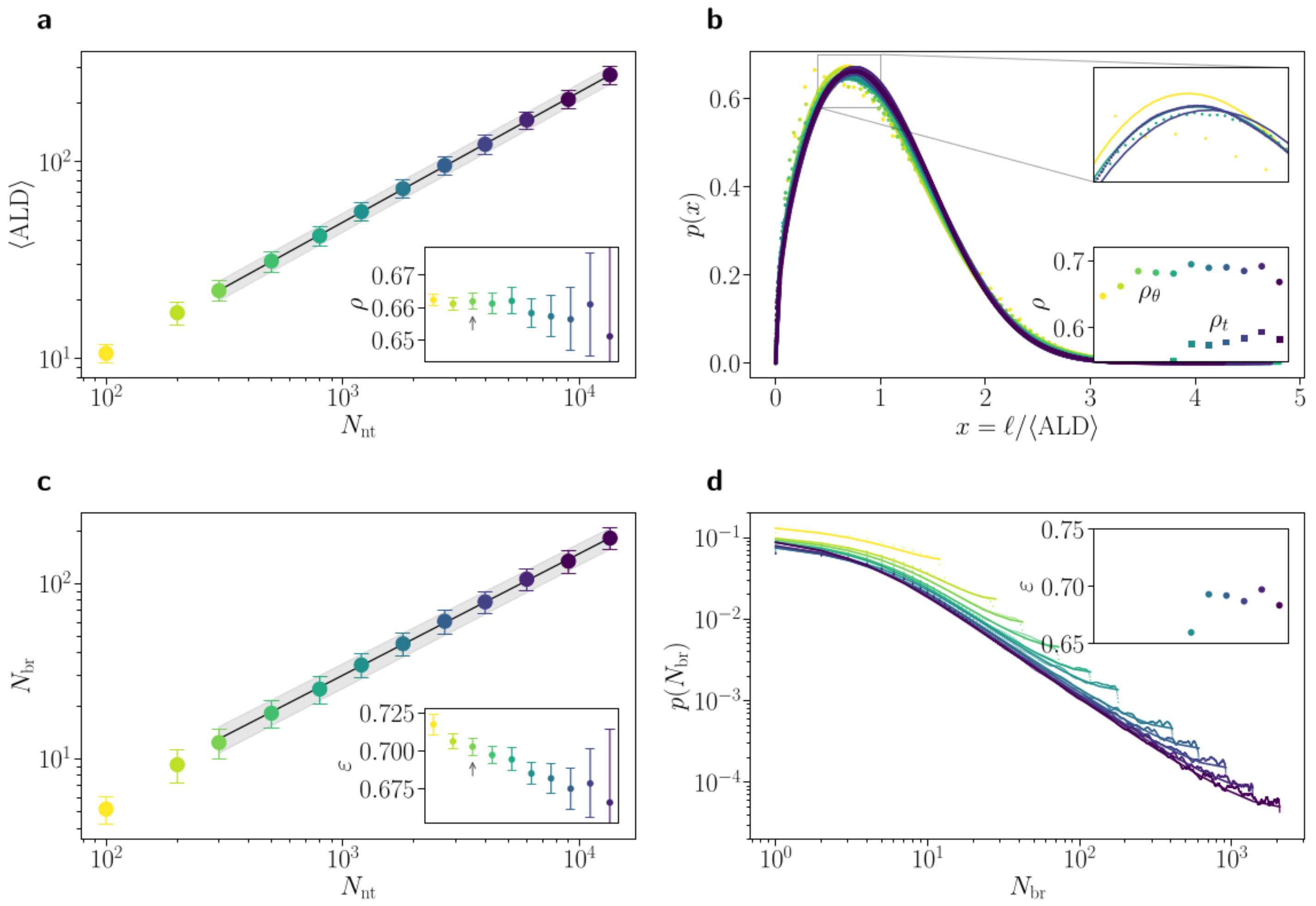
Polymer model (3D)	ν	ρ	ε	ν_{Flory}	ρ_{Flory}
Ideal linear	1/2	1	1	1/2	1
Self-avoiding linear	0.5877	1	1	3/5	1
Ideal branching	1/4	1/2	1/2	1/4	1/2
Self-avoiding branching	1/2	0.654	0.651	7/13	9/13

(the only known exponent for RNA is ρ = 0.67)

RNA as a branched polymer: pipeline

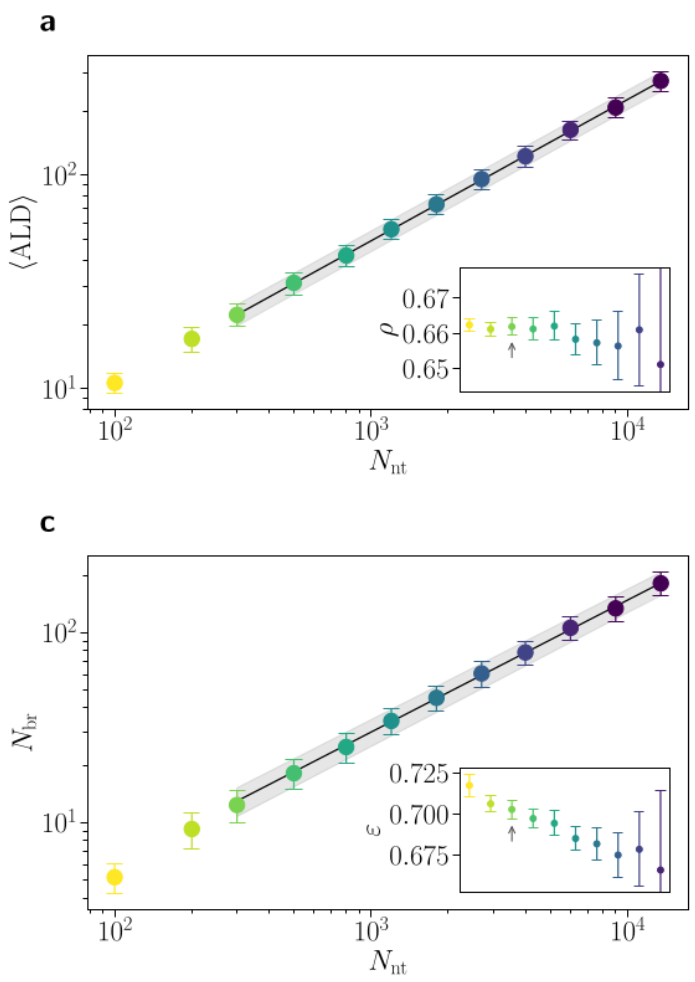


Scaling properties of RNA



- sequence length dependence vs. individual distributions
- distribution of branch sizes → **exponent ε**
- distribution of path lengths → **exponent ρ**

Scaling properties of RNA: sequence length dependence

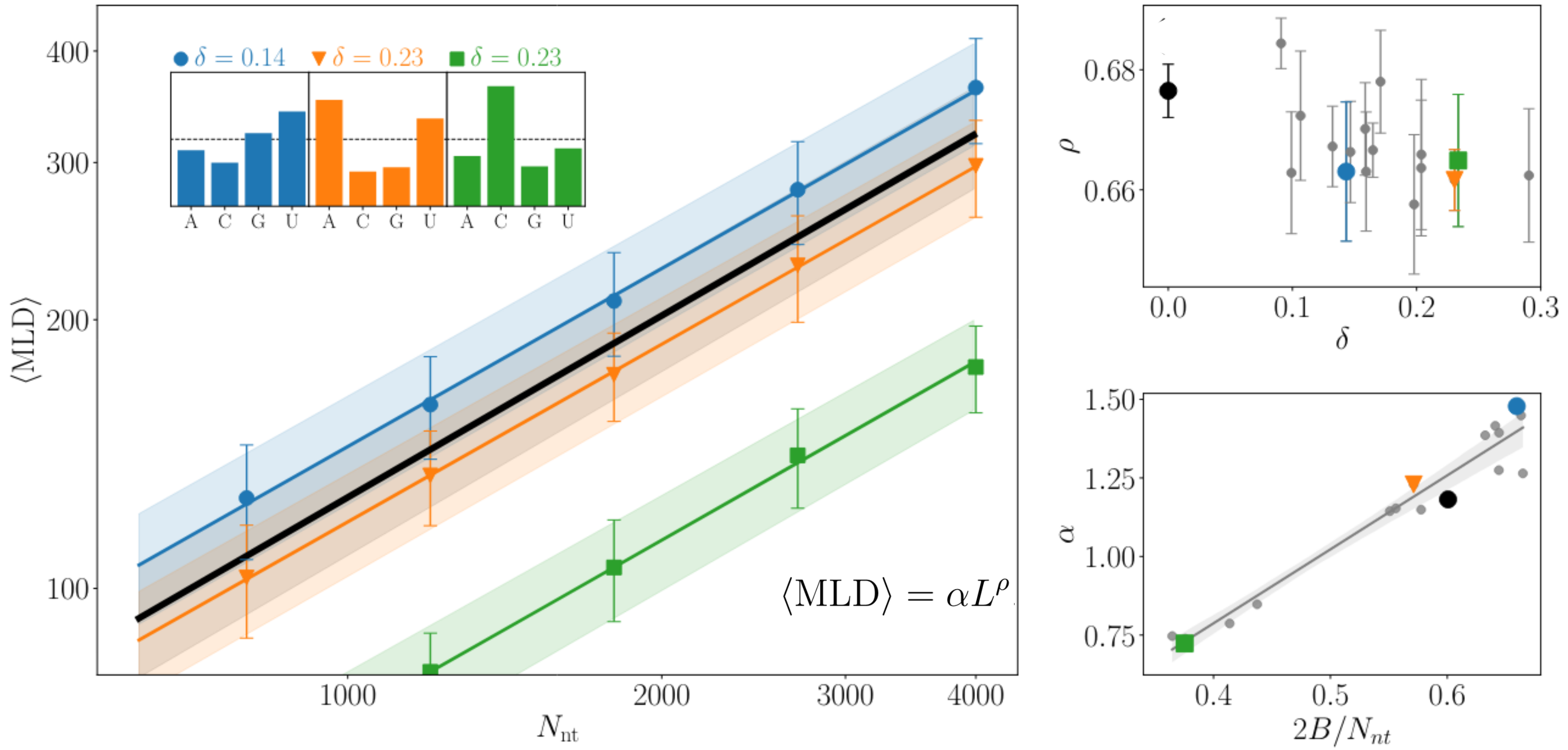


$$\langle \text{ALD} \rangle \sim N_{\text{nt}}^{\rho}$$

$$\langle N_{\text{br}} \rangle \sim N_{\text{nt}}^{\varepsilon}$$

- **sequence length dependence** vs. individual distributions
- distribution of branch sizes → **exponent ε**
- distribution of path lengths → **exponent ρ**

Scaling properties of RNA: nucleotide composition

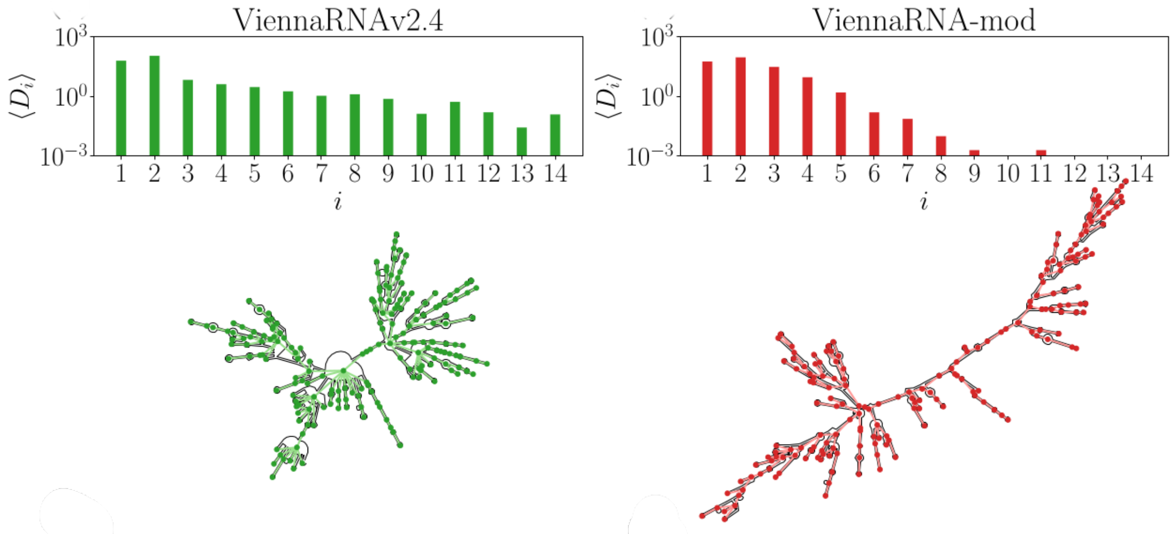


Vaupotič et al, arXiv:2212.00829

- scaling exponents ρ and ε independent of RNA nucleotide composition
- prefactor depends on the amount of base pairs formed

Scaling properties of RNA: multiloop energy parameters

Energy model	E_0	E_{un}	E_{br}
Turner1999	10.1	−0.3	−0.3
Turner2004	9.25	NA	0.63
ViennaRNA (< v2.0)	3.4	0.0	0.4
ViennaRNA (v2.0+)	9.3	0.0	−0.9
RNAstructure	9.3	0.0	−0.6
Andronescu2007	4.4	0.04	0.03
Langdon2018	9.3	0.0	−0.8

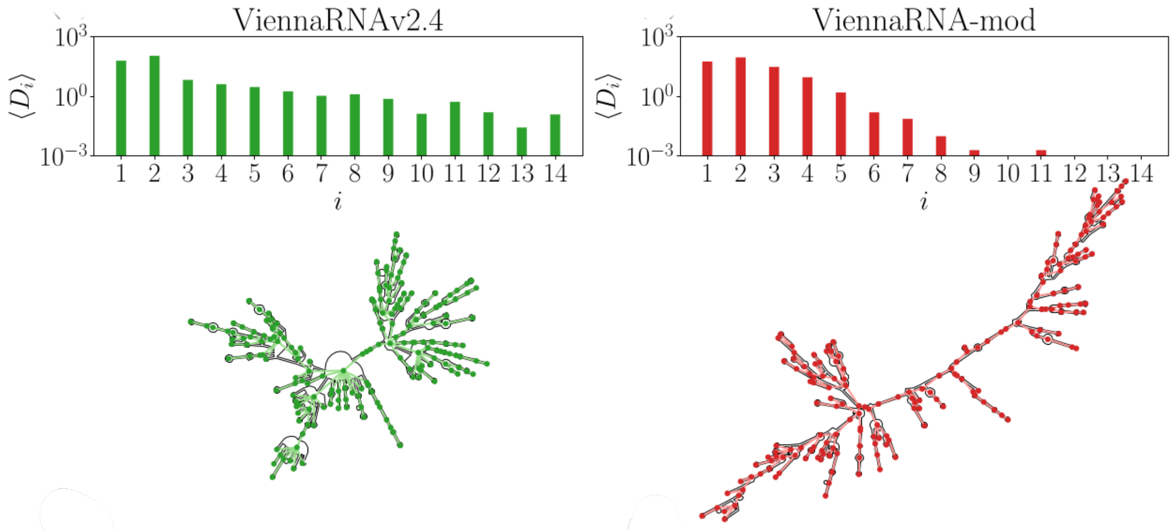


$$E_{\text{multiloop}} = E_0 + E_{\text{br}} \times [\text{branches}] + E_{\text{un}} \times [\text{unpaired nucleotides}]$$

- multiloop energy parameters lead to different node degree distributions and different branching structures of the same RNA sequence

Scaling properties of RNA: multiloop energy parameters

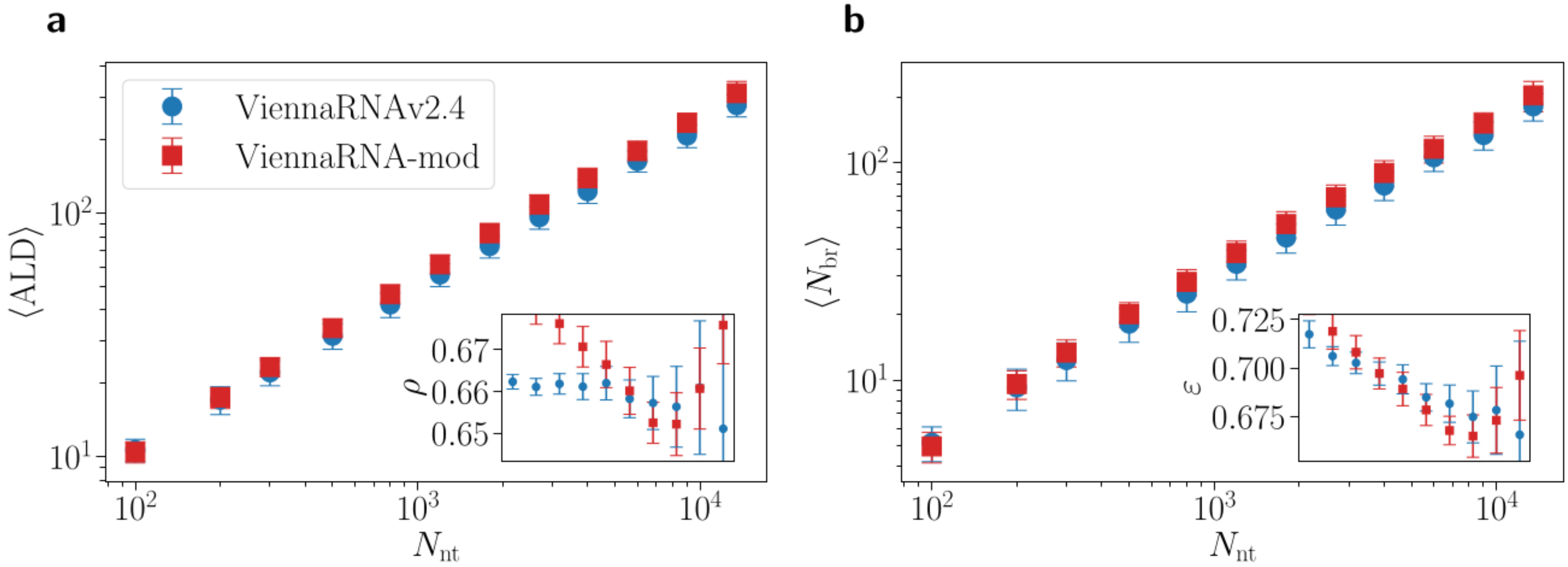
Energy model	E_0	E_{un}	E_{br}
Turner1999	10.1	−0.3	−0.3
Turner2004	9.25	NA	0.63
ViennaRNA (< v2.0)	3.4	0.0	0.4
ViennaRNA (v2.0+)	9.3	0.0	−0.9
RNAstructure	9.3	0.0	−0.6
Andronescu2007	4.4	0.04	0.03
Langdon2018	9.3	0.0	−0.8



$$E_{\text{multiloop}} = E_0 + E_{\text{br}} \times [\text{branches}] + E_{\text{un}} \times [\text{unpaired nucleotides}]$$

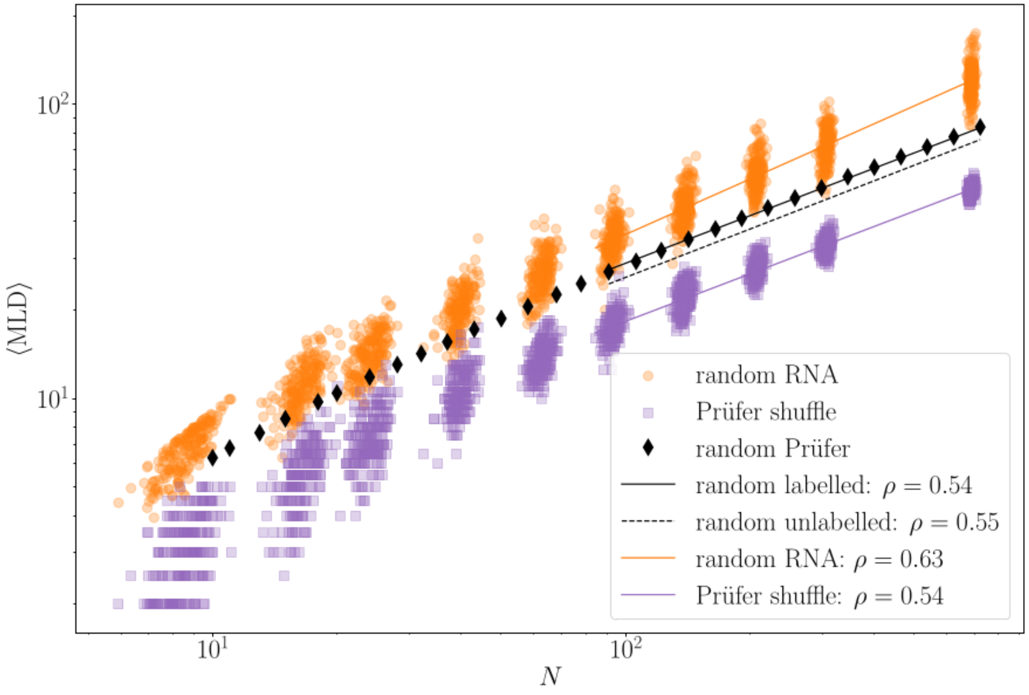
- multiloop energy parameters lead to different node degree distributions and different branching structures of the same RNA sequence

Scaling properties of RNA: multiloop energy parameters

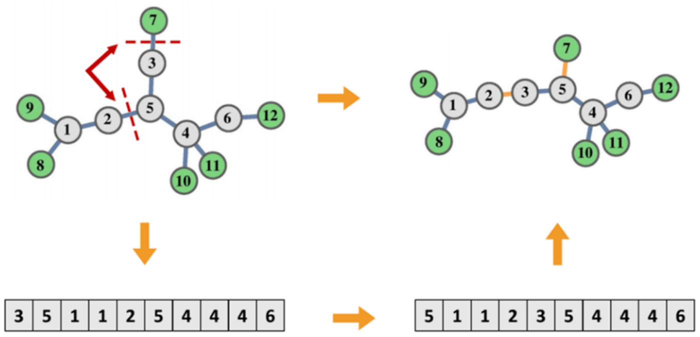


- significantly different (multiloop) energy parameters do not lead to differences in scaling

Scaling properties of RNA: node degree distribution



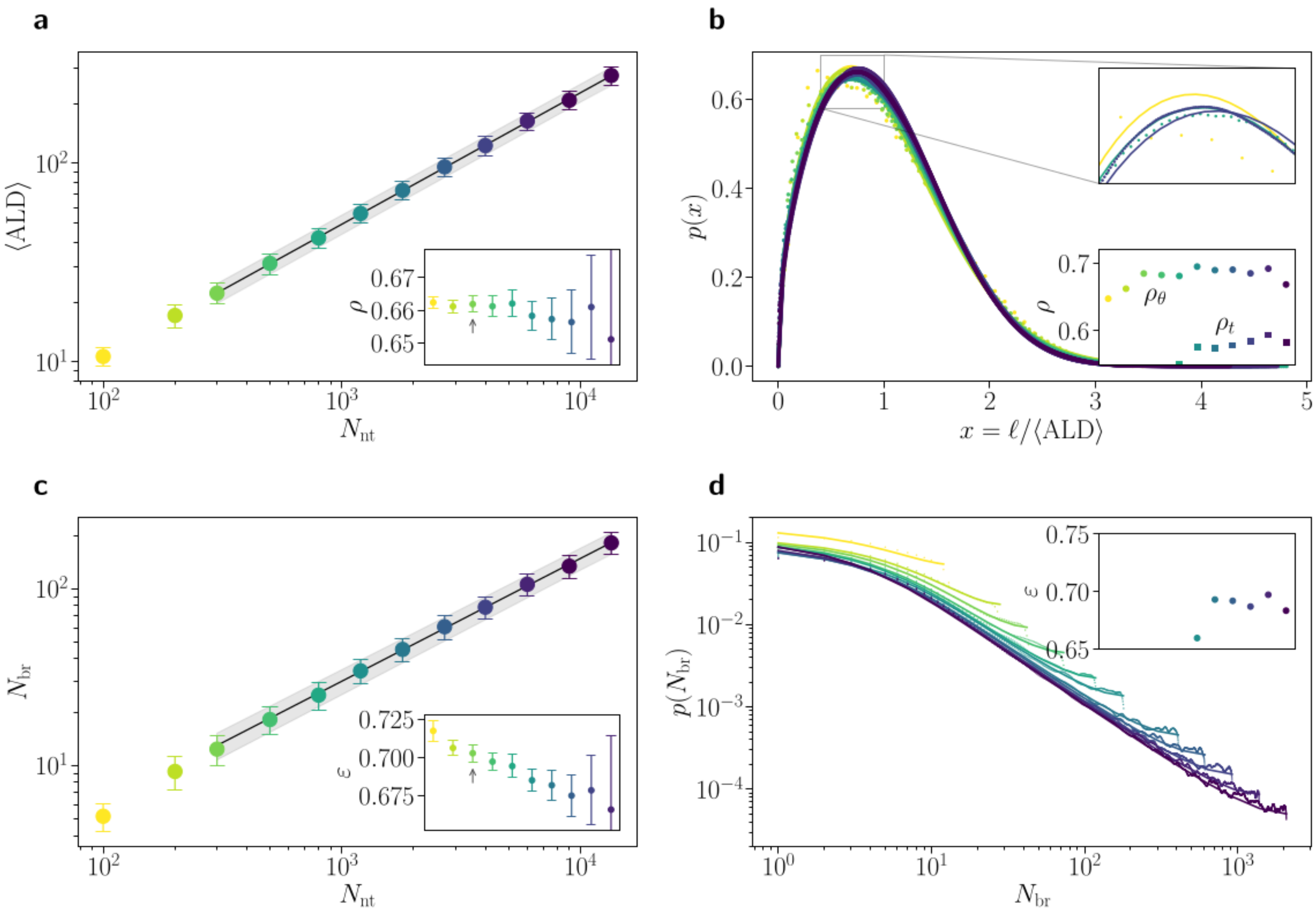
Prüfer shuffle retains the node degree distribution while reshuffling the nodes



SW Singaram et al, J Phys Chem B 2016

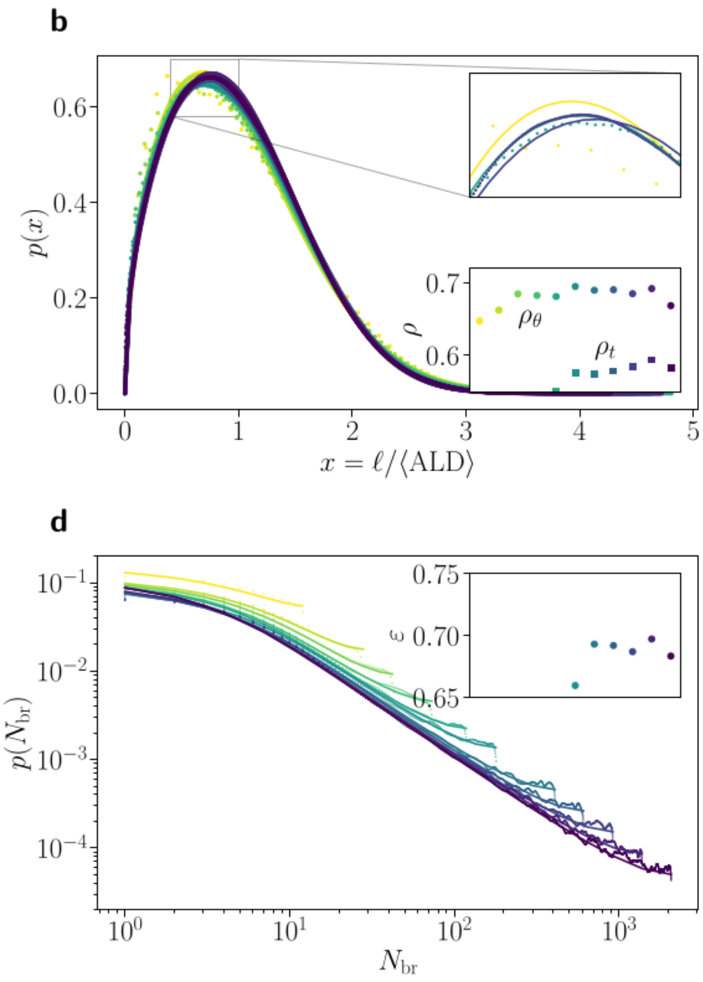
- Prüfer shuffle drastically changes (lowers) the scaling exponents ρ and ε

Scaling properties of RNA



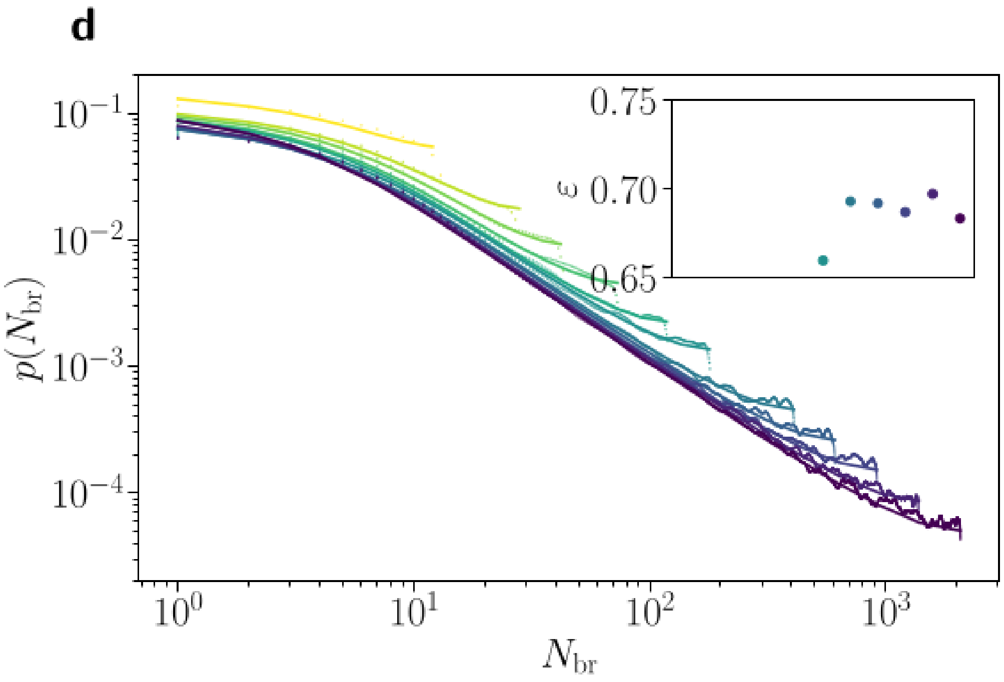
- sequence length dependence vs. individual distributions
- distribution of branch sizes → **exponent ε**
- distribution of path lengths → **exponent ρ**

Scaling properties of RNA: individual distributions



- sequence length dependence vs. **individual distributions**
- distribution of branch sizes → **exponent ε**
- distribution of path lengths → **exponent ρ**

Scaling properties of RNA: individual distributions

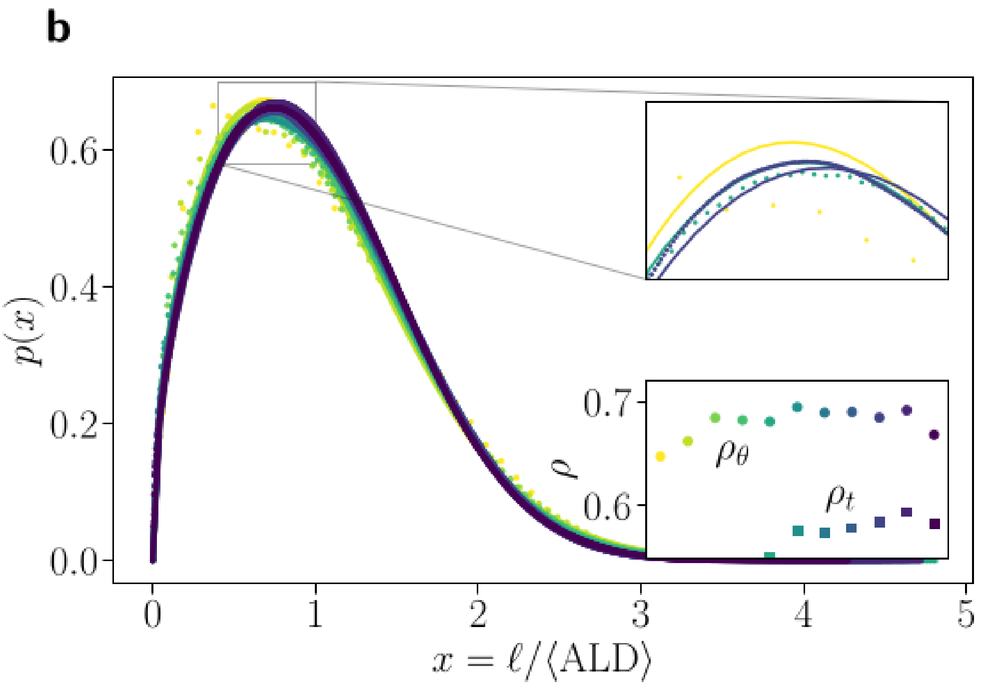


- distribution of branch sizes → exponent ε for RNA of fixed length

$$p(N_{\mathrm{br}}) \sim \left(\frac{N}{N_{\mathrm{br}}(N - N_{\mathrm{br}} - 1)} \right)^{2-\varepsilon}$$

- (thermal+sequence averages improve the data)
- applicable even to biological RNAs whose length variation is negligible

Scaling properties of RNA: individual distributions



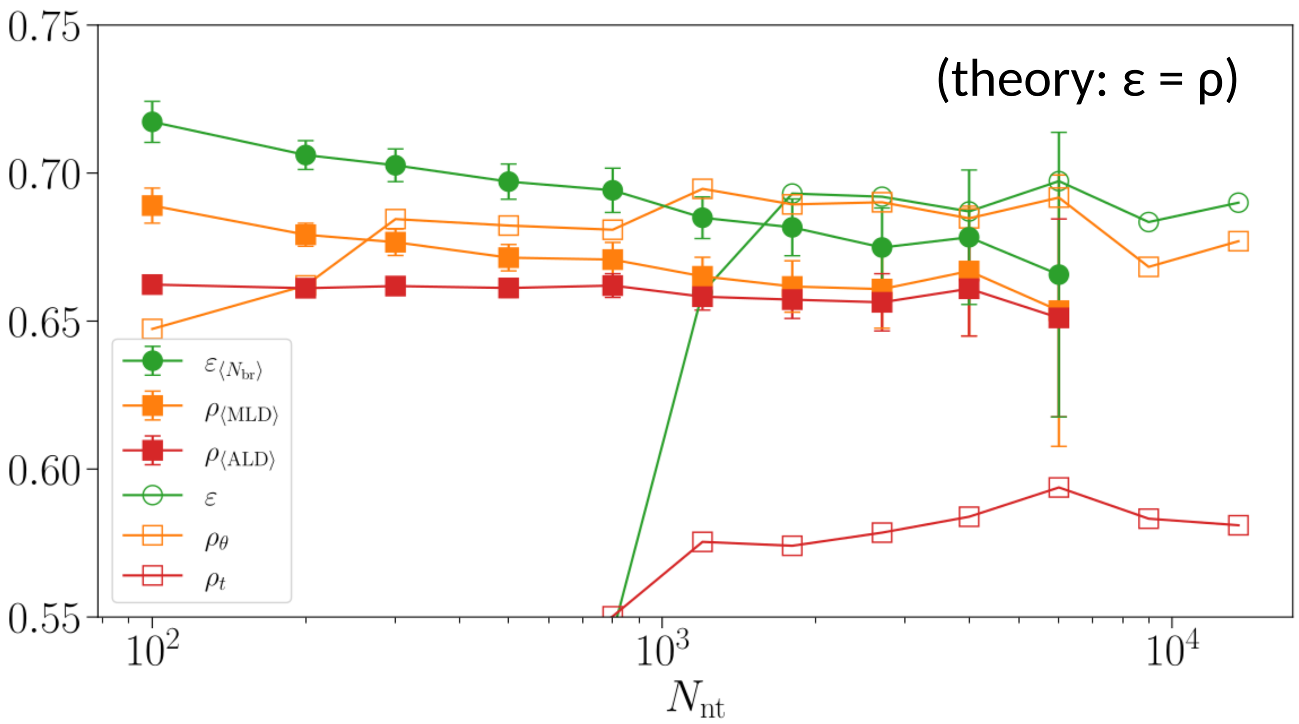
- distribution of path lengths → exponent ρ for RNA of fixed length
- two-parametric distribution (Redner-des-Cloizeaux) → two different ways of obtaining ρ!

$$p(\ell) = \frac{1}{\langle \mathrm{ALD} \rangle} \, q\left(\frac{\ell}{\langle \mathrm{ALD} \rangle} \right)$$

with

$$q(x) = C \, x^{\theta} \exp\left(-(Kx)^{t} \right)$$

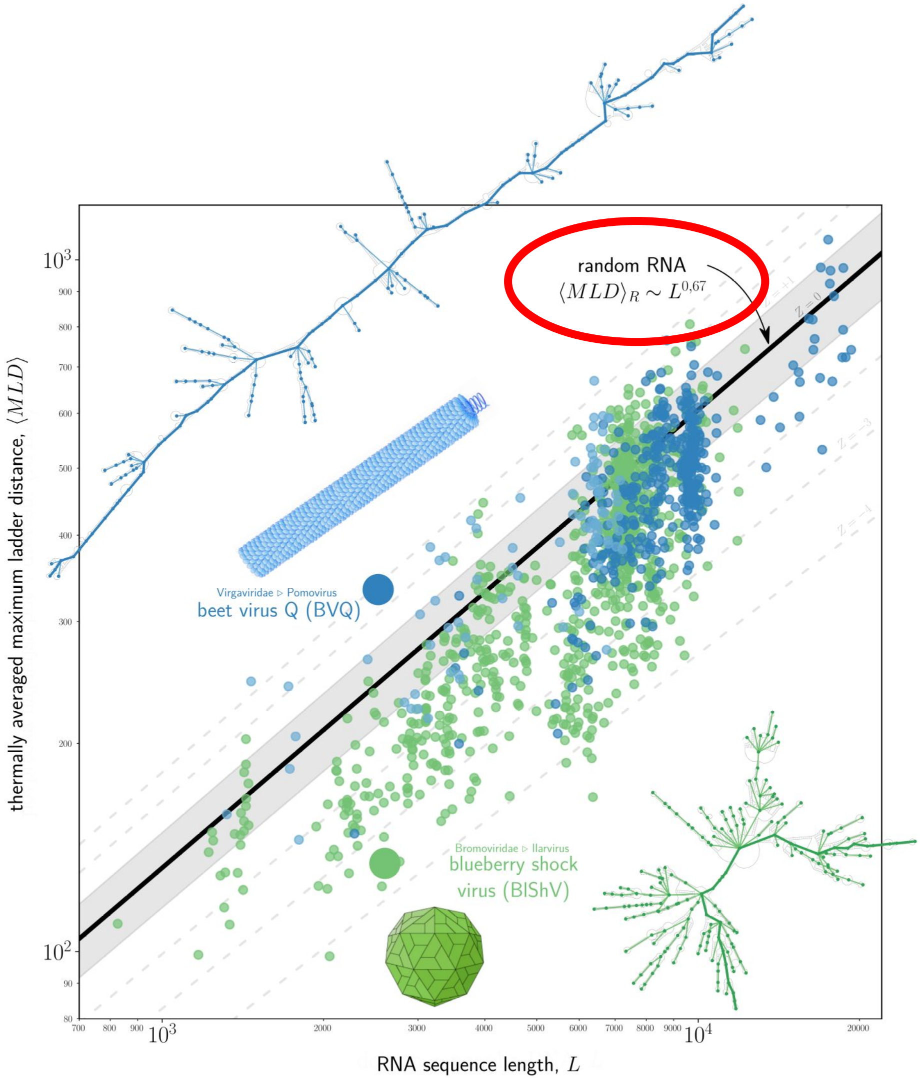
RNA as a randomly branched polymer



- exponents obtained from scaling match each other more closely
- discrepancy between exponents ρ obtained from branch weight distribution

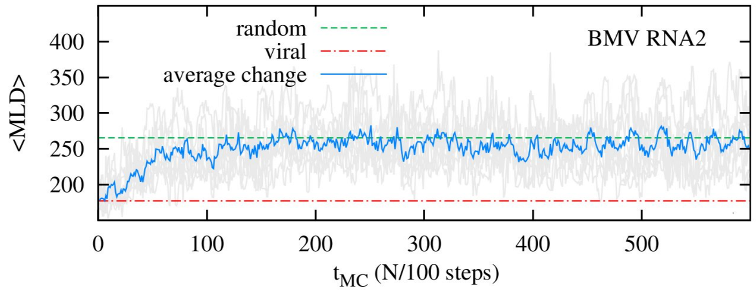
Branching structure of RNA

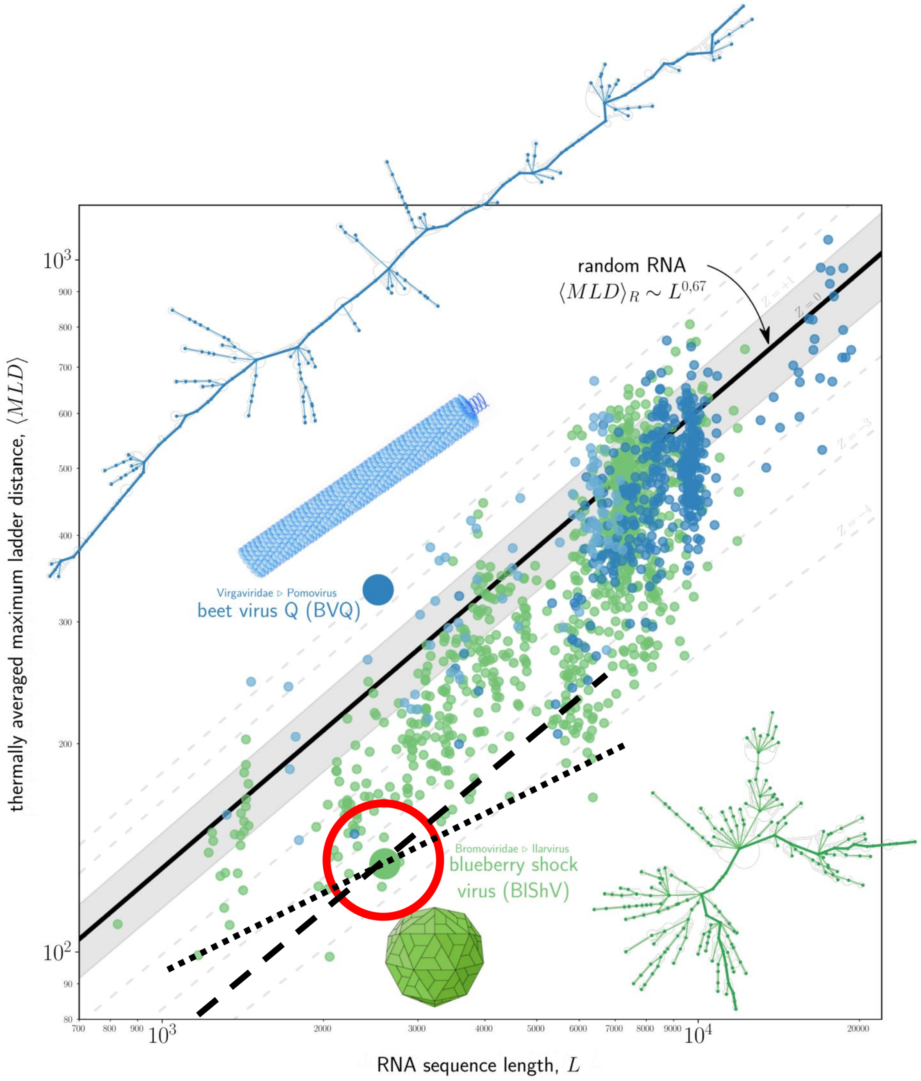
- RNA as a randomly branched polymer: scaling exponents ε and ρ close to those of branched SAW polymers
- distributions of path lengths and branch weights → applicable to RNAs of fixed length
- what makes the branching properties of (some) RNA viruses different?



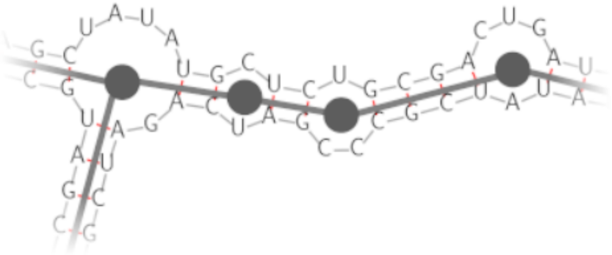
Branching structure of *viral* RNA

- RNA as a randomly branched polymer: scaling exponents ε and ρ close to those of branched SAW polymers
- distributions of path lengths and branch weights → applicable to RNAs of fixed length
- what makes the branching properties of (some) RNA viruses different?



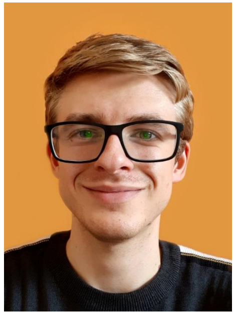


L Tubiana et al, Biophys J 108 2015



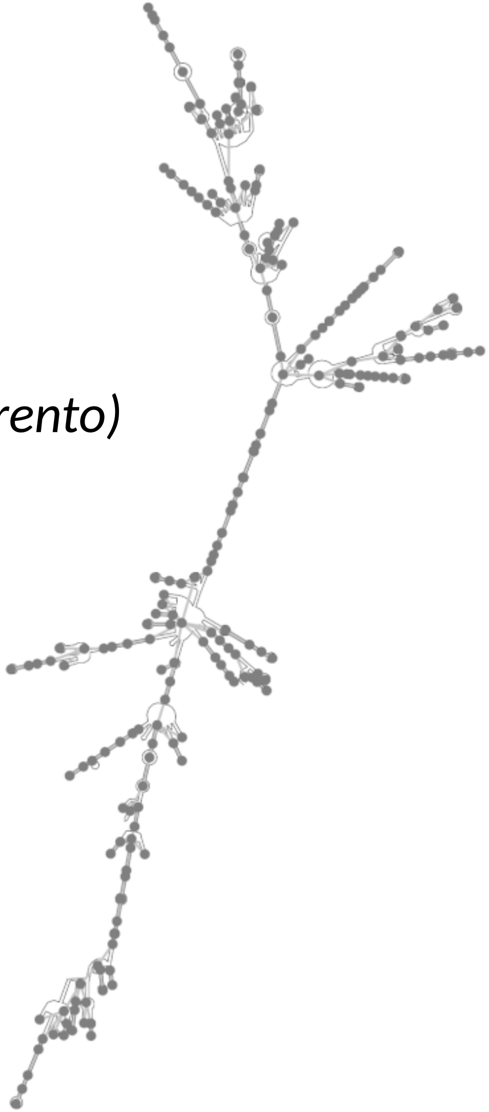
Acknowledgments

Domen Vaupotič (IJS)



Luca Tubiana (Uni. Trento)

Angelo Rosa (SISSA)



Thank you for your attention!