

# GENOME PACKING AND STABILITY OF VIRAL CAPSIDS



**Jure Dobnikar**

*Institute of Physics  
Chinese Academy of Sciences, Beijing*

## Who did the work

James D. Farrell



Tine Curk



Jure Dobnikar



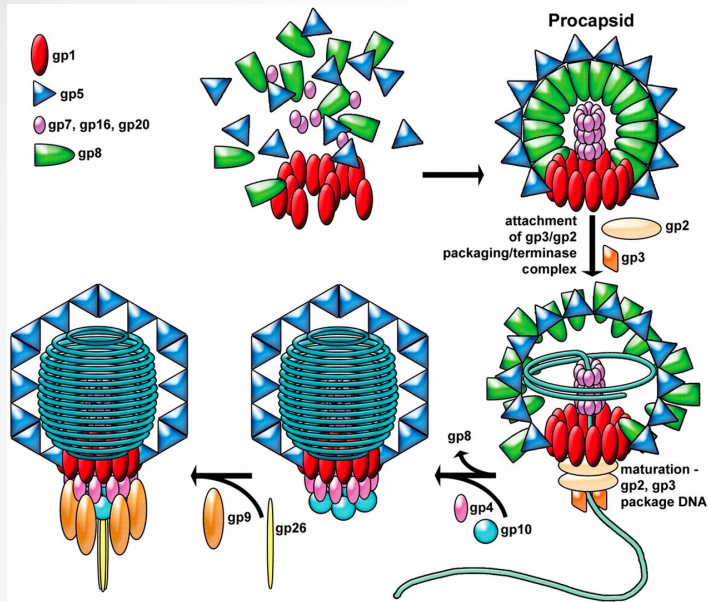
Rudi Podgornik



## Who paid for it

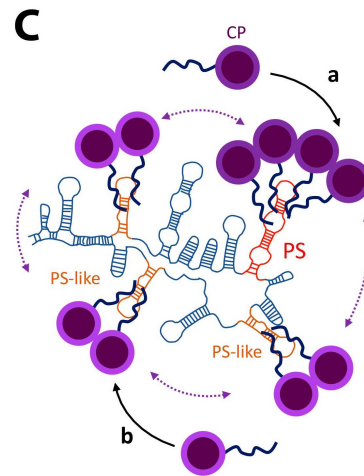
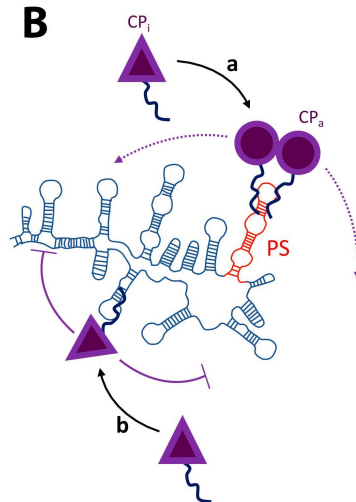
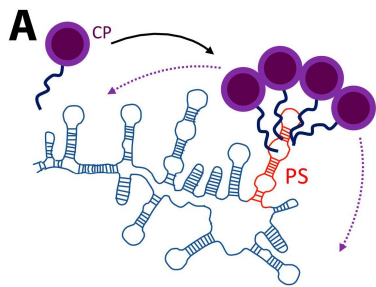
Key project (12034019) of the National Natural Science Foundation of China  
Chinese National Science Foundation (grants 11874398, 12034019)  
Strategic Priority Research Program of CAS (XDB33000000)  
International collaboration grant from K. C. Wong Educational Foundation.

# Genome packing in viruses



## Pre-assembled capsid

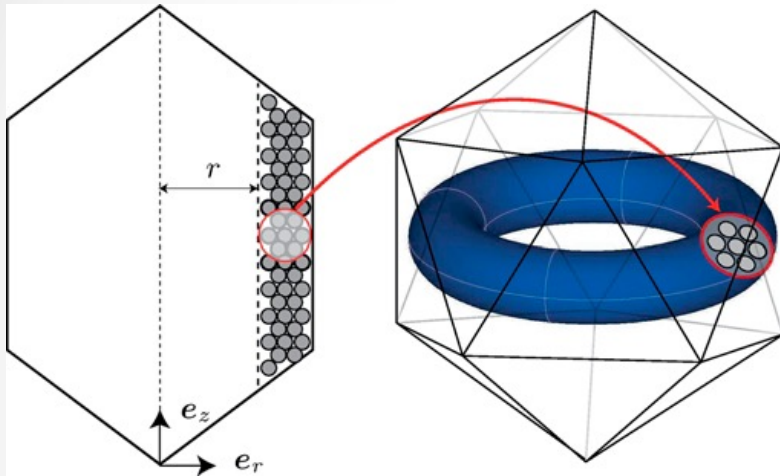
- motor-assisted packing (typical for dsDNA)
- under large pressure when loaded



## Co-assembled capsid

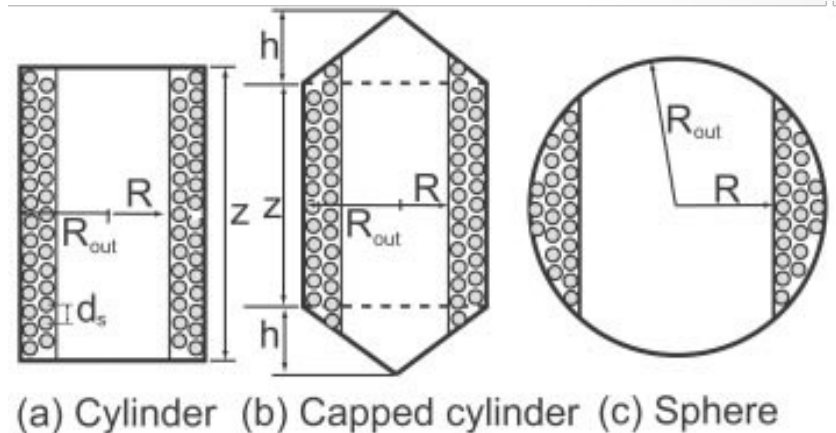
- RNA with PS as template

# Packing dsDNA: spool conjecture



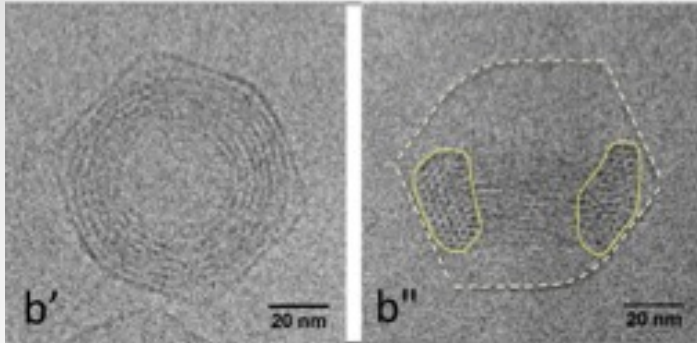
W.S. Klug, M.T. Feldman, M. Ortiz:  
*Three-dimensional director-field predictions  
of viral DNA packing arrangements,*  
Computational Mechanics **35** 146 (2004)

P.K. Purohit, J. Kondev, R.Philips:  
*Mechanics of DNA packing in viruses,*  
PNAS **100** 3173 (2003)

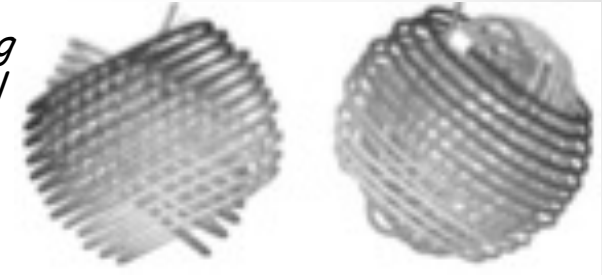




# More complex structures?



Petrov, Anton S., and Stephen C. Harvey, *Packaging double-helical DNA into viral capsids: structures, forces, and energetics*, *Biophysical journal* 95 497 (2008)



Leforestier, Amélie, et al., *Protein-DNA interactions determine the shapes of DNA toroids condensed in virus capsids*, *Biophysical journal* 100 2209 (2011)

## Packing of elastic wires in spherical cavities

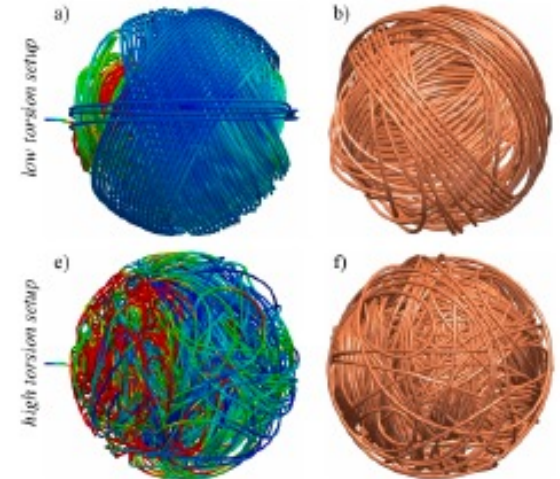
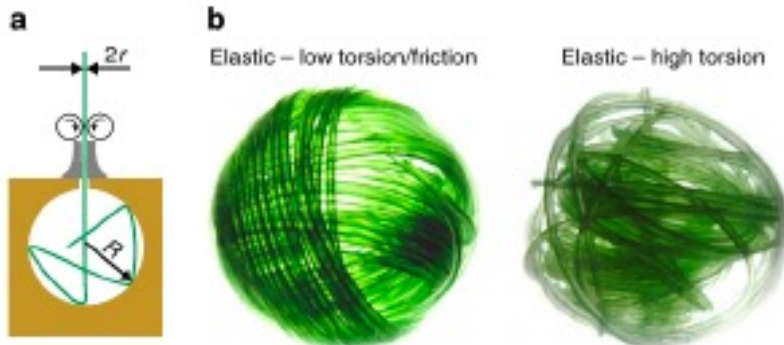
N. Stoop<sup>1</sup>, J. Najafi<sup>2</sup>, F. K. Wittel<sup>1</sup>, M. Habibi<sup>2</sup>, and H. J. Herrmann<sup>1,3</sup>

<sup>1</sup> Computational Physics for Engineering Materials, ETH Zurich, Schafnattstr. 6, HIF, CH-8093 Zurich, Switzerland

<sup>2</sup> Department of Physics, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan 45137-66731, Iran

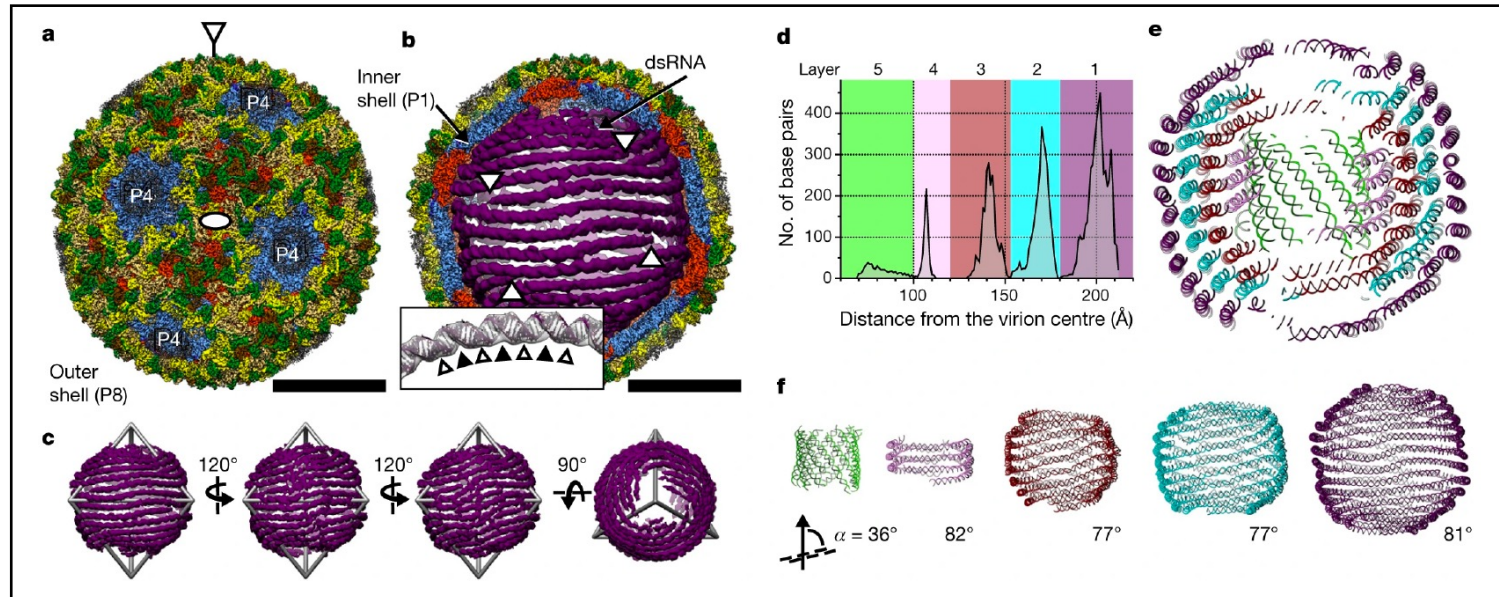
<sup>3</sup> Departamento de Física, Universidade Federal do Ceará, Campus do Pici, 60451-970 Fortaleza, Ceará, Brazil

(Dated: February 4, 2011)



R. Shaebani, J. Najafi, A. Farnudi, D. Bonn, M. Habibi: *Compaction of quasi-one-dimensional elastoplastic materials*, *Nature Communications* 8 15568 (2017)

# Spool - twisted nematic



## Cryo-EM of bacteriophage $\phi 6$

- Bacteriophage  $\phi 6$  adopts multiple spool domains around the portal axis; the interior is unresolved<sup>a</sup>
- Since dsRNA can kink, we can imagine nested spools wrapped around a perpendicular, (twisted) nematic domain

<sup>a</sup>Serban L Ilca et al. "Multiple liquid crystalline geometries of highly compacted nucleic acid in a dsRNA virus". In: *Nature* 570.7760 (2019), pp. 252–256.

# Questions

- **Ground state** configurations of elastic filaments in spherical confinement ?  
*(coarse grained model of DNA or metallic wires...)*
- Effects of nonlinear elasticity (**kinks**) and **kinetics** on DNA packing in viral capsid

# Model

Worm-like chain (length  $L$ , cross-section  $\sigma$ ) with excluded volume persistence length  $l_p$  (elastic modulus  $K_c$ ), Spherical container radius  $R_0$

Limit: long, thin, rigid polymer:  $R_0/L \sim 0$ ,  $\sigma/R_0^2 \sim 0$ ,  $R_0/l_p \sim 0$

$$E_{el} = \frac{1}{2} K_c \int_0^L ds / R^2(s)$$

## Continuum theory of polymer nematics

polymer “current”  $\mathbf{t}(\mathbf{r})$   $(\nabla \cdot \mathbf{t} = 0)$

$$E_{el} = \frac{2K_c}{\pi\sigma^2} \int_V t (\hat{\mathbf{t}} \times (\nabla \times \hat{\mathbf{t}}))^2 d\mathbf{r},$$

$$V_p = \int_C t d\mathbf{r}, \text{ with } t \equiv |\mathbf{t}|$$

$$t^m = (\pi\sqrt{3}/6) \approx 0.907 \quad \text{local hexagonal packing}$$

$$\nu \equiv V_p / (t_m V_0) = (3\sqrt{3}\sigma^2 / 8\pi R_0^3) L$$

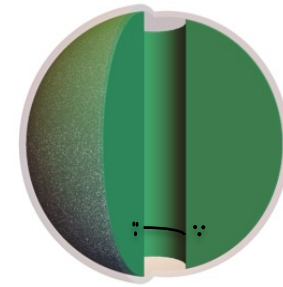


## Single spool:

$$\hat{\mathbf{t}}^{1s} = t_m(-y, x, 0) / \sqrt{x^2 + y^2},$$

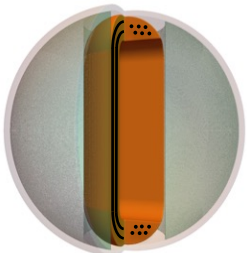
$$E_{el}^{1s}(\nu, R_0) = \frac{\pi^2 KR_0}{\sqrt{3}} \log \left( -\nu^{1/3} + \frac{1}{2} \ln \frac{1 + \nu^{1/3}}{1 - \nu^{1/3}} \right)$$

diverging elastic penalty with dense packing

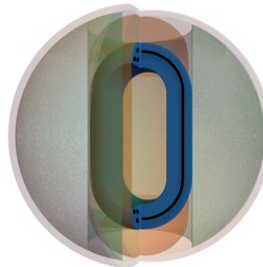


## Other options for material distribution:

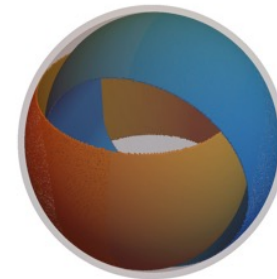
double spool



triple spool



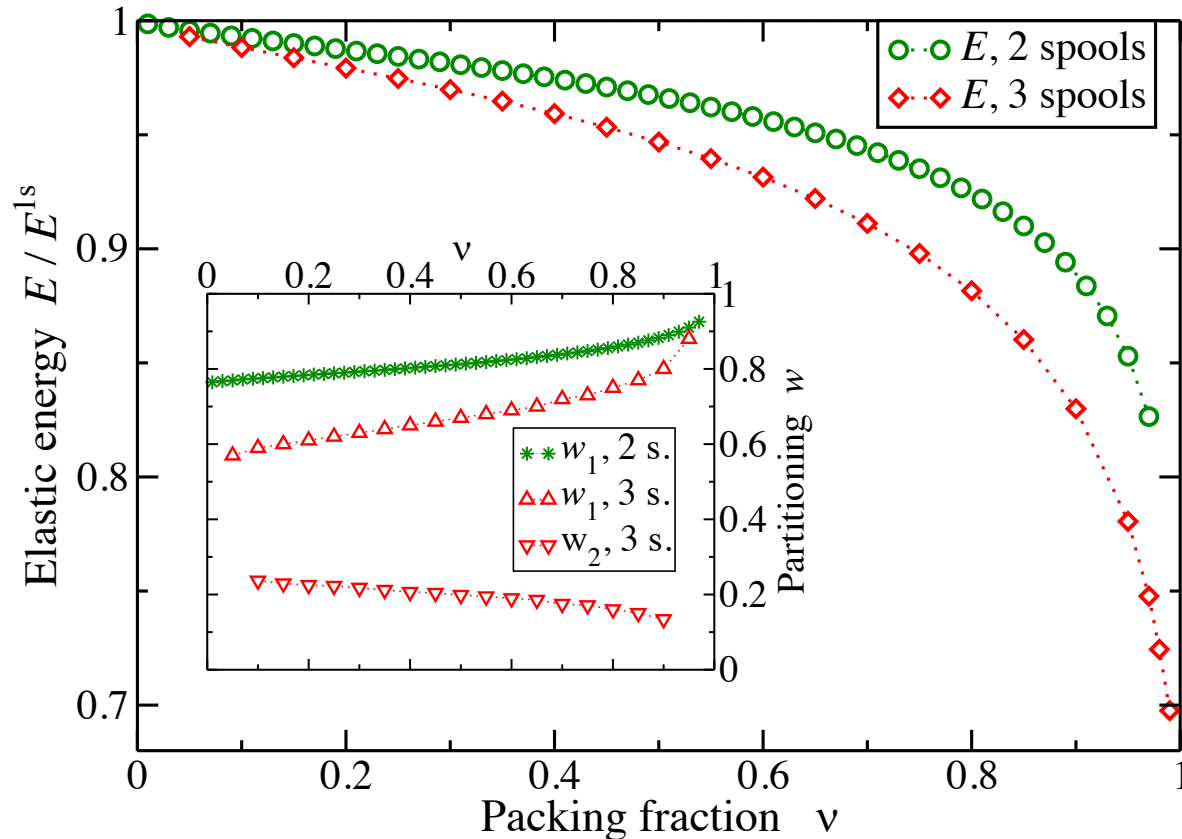
Hopf link



$$E_{el}^{2s}(\nu_1, \nu_2, R_0) = E_{el}^{1s}(\nu_1, R_0) + E_{el}^{1s}(\nu_2^t, R_1)$$



# Multidomain solutions always better !



# Simulation: Semiflexible Polymer

harmonic bonds;  
 cosine angles;  
 WCA non-bonded  
 WCA wall-polymer

$$V_{12} = k_b (r - r_{eq})^2$$

$$V_{13} = k_\theta (1 + \cos \theta)$$

$$V_{nb} = \begin{cases} 4\epsilon \left[ \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right] + V_{shift} & r \leq 2\frac{1}{6}\sigma \\ 0 & r > 2\frac{1}{6}\sigma \end{cases}$$

parallel tempering MD with 336  
 temperature replicas  
 $(T_{lo}, T_{hi}) = (1 T^*, 300 T^*)$

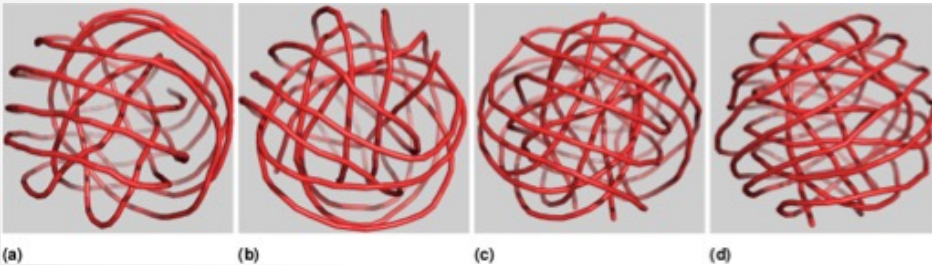
$$r_{eq} = \sigma$$

$$k_b = 5000\epsilon / \sigma^2$$

$$k_\theta = 300\epsilon$$

$$R = 5\sigma$$

## Low energy snapshots

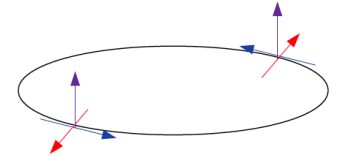


Energy minimization  
 Clustering / Domain decomposition algorithm

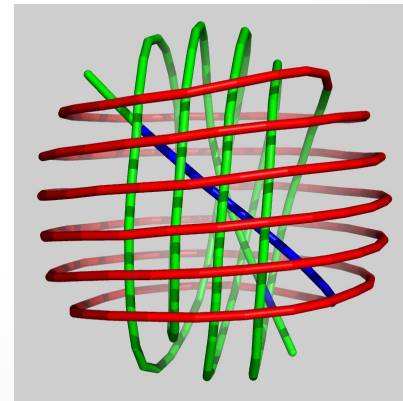
## Binormal Vectors

rather than bond vectors, we  
 will look at binormal vectors

for a perfect spool, the  
 binormal vectors are identical  
 at all points on the chain



$$\hat{\mathbf{B}} = \hat{\mathbf{T}} \times \hat{\mathbf{N}} = \frac{\mathbf{r}' \times \mathbf{r}''}{|\mathbf{r}' \times \mathbf{r}''|}$$



# Bond Vectors

repurpose the bond-orientational order parameters of Steinhardt, Nelson, and Ronchetti

designed to quantify cluster symmetry based on interparticle vectors (*bond vectors*)

$$Q_{lm}(\vec{r}) = Y_{lm}(\theta(\vec{r}), \phi(\vec{r}))$$

$$\bar{Q}_{lm_1} = \langle Q_{lm}(\vec{r}) \rangle$$

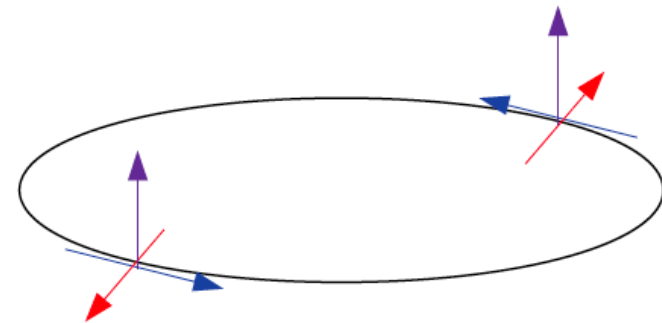
$$Q_l = \left( \frac{4\pi}{2l+1} \sum_{m=-l}^l |\bar{Q}_{lm}|^2 \right)^{1/2}$$

$$W_l = \sum_{\substack{m_1, m_2, m_3 \\ m_1 + m_2 + m_3 = 0}} \begin{pmatrix} l & l & l \\ m_1 & m_2 & m_3 \end{pmatrix} \bar{Q}_{lm_1} \bar{Q}_{lm_2} \bar{Q}_{lm_3}$$

# Binormal Vectors

rather than bond vectors, we will look at binormal vectors

for a perfect spool, the binormal vectors are identical at all points on the chain



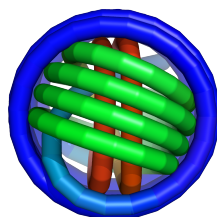
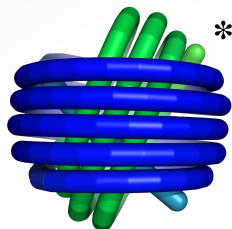
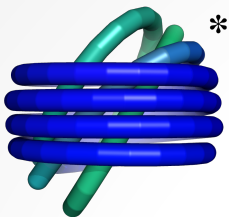
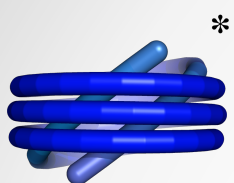
$$\hat{\mathbf{B}} = \hat{\mathbf{T}} \times \hat{\mathbf{N}} = \frac{\mathbf{r}' \times \mathbf{r}''}{|\mathbf{r}' \times \mathbf{r}''|}$$

N=100

150

200

250



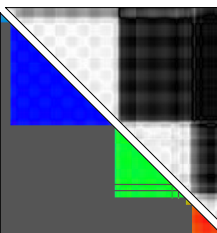
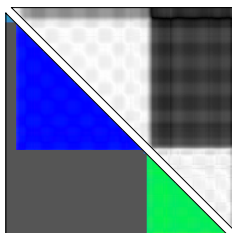
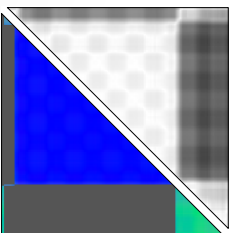
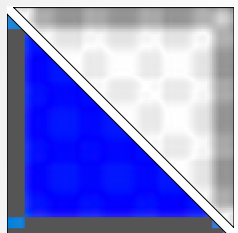
(a)

(b)

(c)

(d)

Triple spool NOT lowest in energy

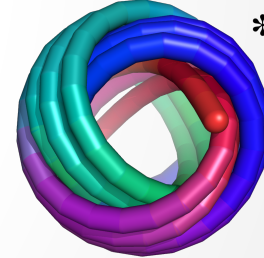
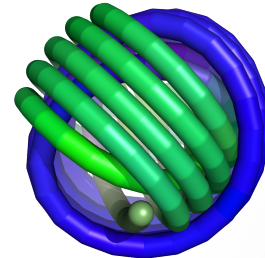
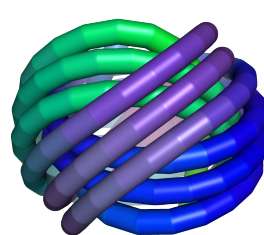
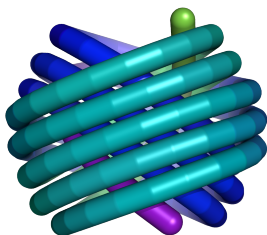


N=200

200

250

250



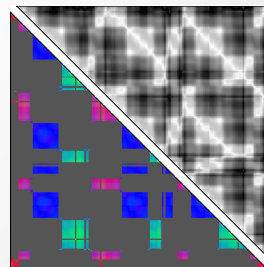
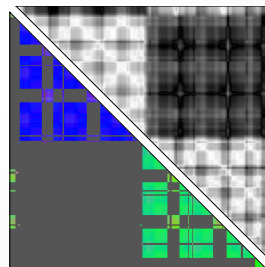
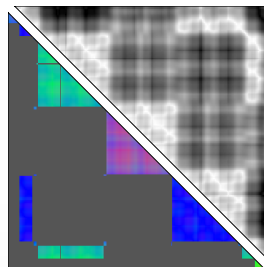
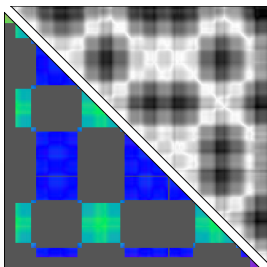
Linked motifs

(a)

(b)

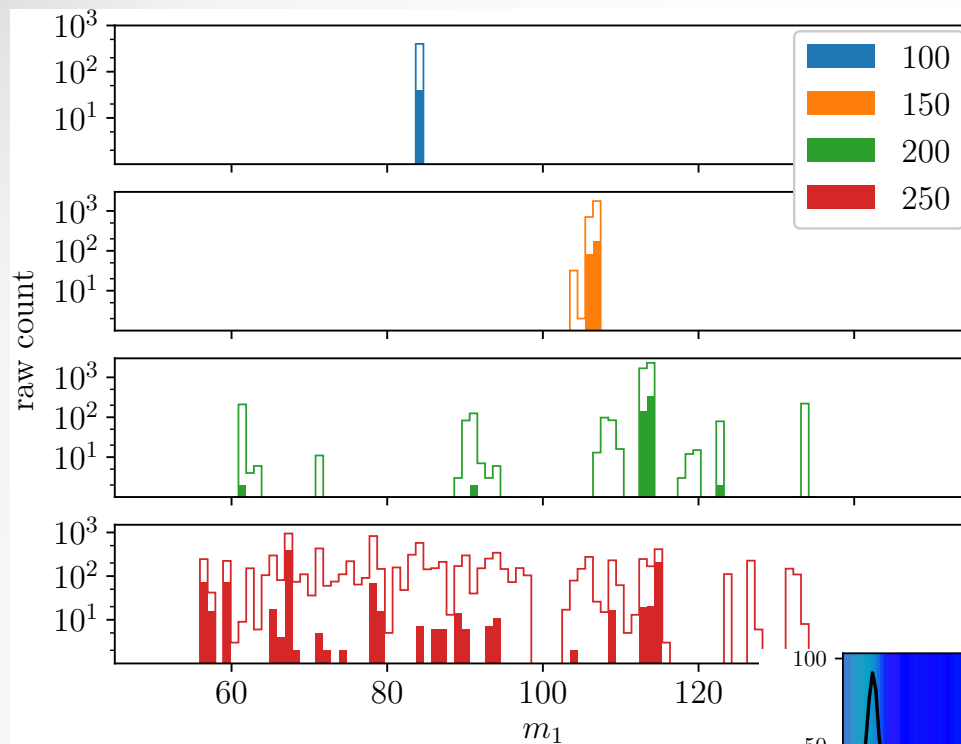
(c)

(d)

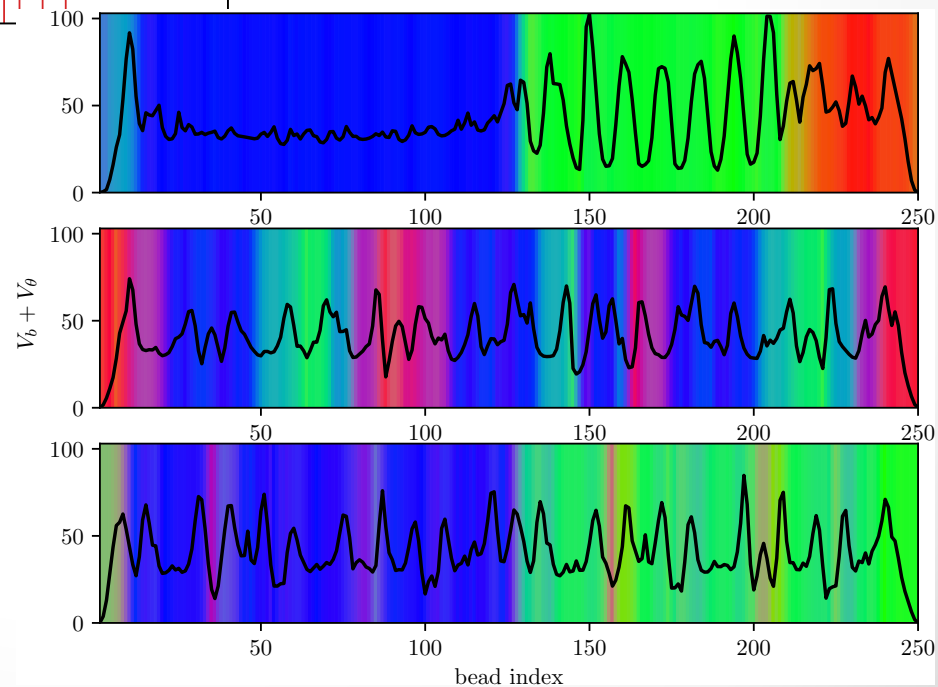




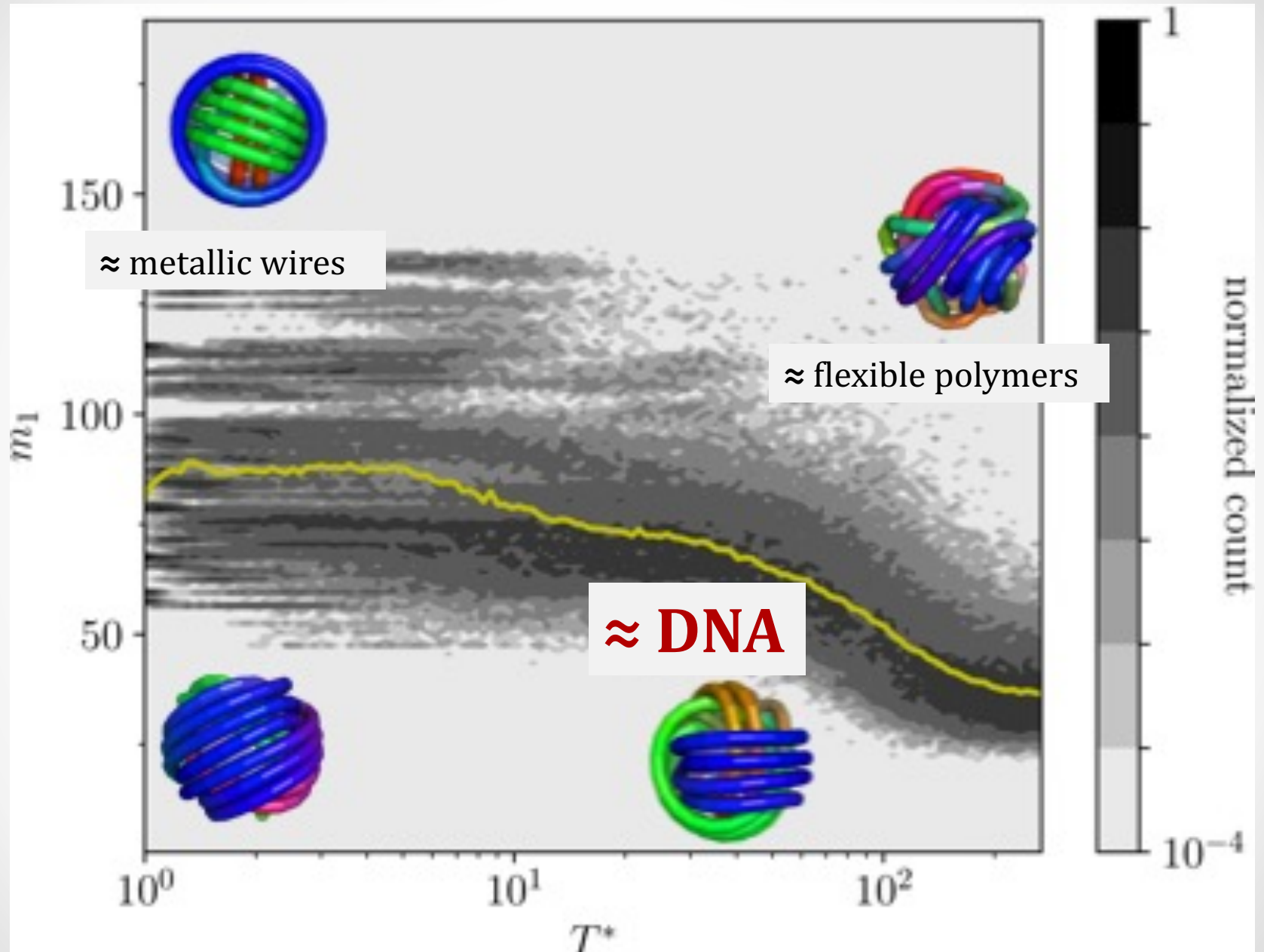
# distribution of mass in 1<sup>st</sup> domain



## Local energy fluctuations



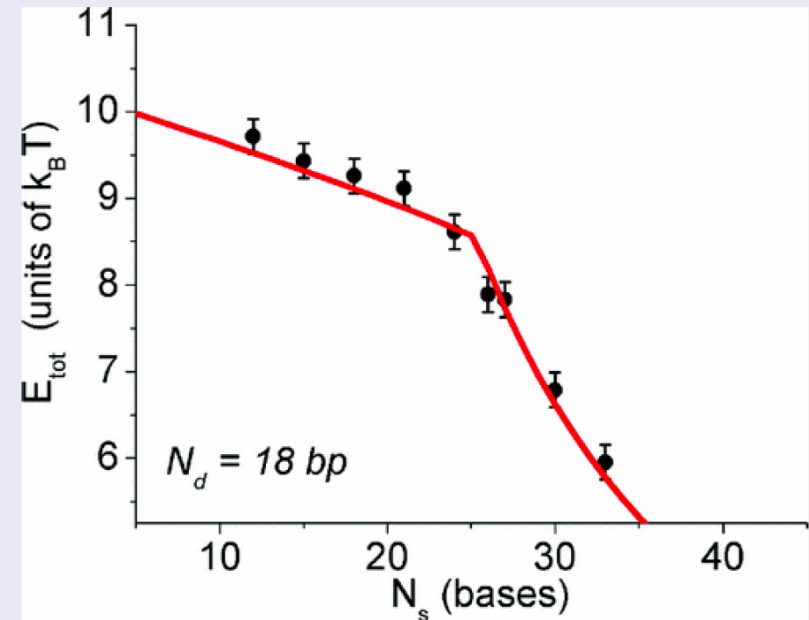
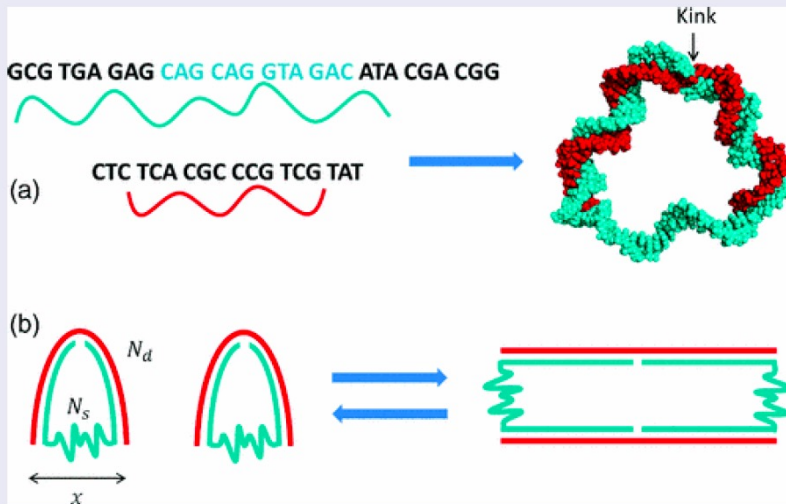
$T^* l_p = \text{const.}$  changing simulation temperature = changing persistence length



Curk, Farrell, Dobnikar, Podgornik, *PRL* (2019)

# Nonlinear elasticity of DNA ... (in progress)

## dsDNA 'kinks'



Upon applying a critical torque ( $\tau_c \approx 31 \text{ pN} \times \text{nm}$ ), dsDNA 'kinks'<sup>a</sup>

<sup>a</sup>Hao Qu et al. "Critical torque for kink formation in double-stranded DNA". In: *PRX* 1.2 (2011), p. 021008.

# DNA with "kinks" ... (in progress)

To model real viruses, we need to make some changes

- dsDNA kinks—the WLC model doesn't take into account non-linear elasticity
- dsDNA is actively packed at finite temperature into the capsid—packed configurations will differ from low-temperature equilibrium structures

Summary of model changes

- 1 WLC  $\xrightarrow{\text{kinks}}$  MWLC
- 2 REMD  $\xrightarrow{\text{driving force}}$  Non-equilibrium MD

A two state dsDNA model—MWLC

Each bead is either hybridised ( $m = 0$ ) or molten ( $m = 1$ ),

$$U^{MWLC}(m, \theta, k_0, k_1, \mu) = \delta_{m,0} U(\theta, k_0) + \delta_{m,1} (U(\theta, k_1) + \mu)$$

where

is the WLC angle potential and

<sup>a</sup>Jie Yan and John F Marko. "Localized single-strand

Adapting the MWLC model for MD simulations

To get the force we need the potential of mean force, which is the free energy of the two states,

$$Q^{MWLC}(\theta, T) = e^{-\beta U(\theta, k_0)} + e^{-\beta(k_1 U(\theta, k_1) + \mu)}$$

$$F^{pmf}(\theta) = -k_B T \log(Q)$$

The force is thus,

$$\begin{aligned} f^{MWLC}(\theta) &= -\frac{\partial F^{pmf}(\theta)}{\partial \theta} \\ &= \sin(\theta) \frac{k_0 e^{-\beta U(\theta, k_0)} + k_1 e^{-\beta(U(\theta, k_1) + \mu)}}{e^{-\beta U(\theta, k_0)} + e^{-\beta(U(\theta, k_1) + \mu)}} \end{aligned}$$

# DNA with "kinks" packed by molecular motor... (in progress)

## dsDNA/dsRNA genomes are packed by molecular motors

- The most powerful motors generate forces as high as **80–100 pN**, and pack to **near crystalline density**<sup>a</sup>
- the fastest known motor, belonging to the T4 bacteriophage, packs at a rate up to  $\approx 2$  kb/s
- complete packaging of the 170-kb T4 genome takes about seven minutes.
- the average rate is  $\approx 700$ bp/s
- bacteriophage  $\phi 6$  has a 13,385 bp genome; at the same packing rate, it would pack in 20 seconds

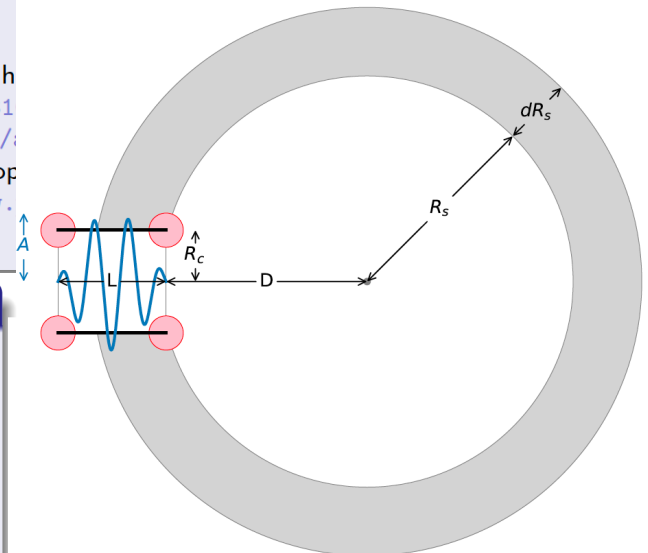
<sup>a</sup>Derek N. Fuller et al. "Single phage T4 DNA packaging motors exhibit large force generation, *h the National Academy of Sciences* 104.43 (2007), pp. 16868–16873. DOI: 10.1073/pnas.0704008104. <https://www.pnas.org/doi/pdf/10.1073/pnas.0704008104>. URL: <https://www.pnas.org/doi/10.1073/pnas.0704008104>. "Single-molecule packaging initiation in real time by a viral DNA packaging machine from bacteriophage  $\phi 6$ , *Science* 324.5936 (2009), pp. 155–158. DOI: 10.1126/science.1170000. eprint: <https://www.pnas.org/doi/abs/10.1073/pnas.1407235111>.

### Driven Packing

- spherical enclosure
- time-dependent force

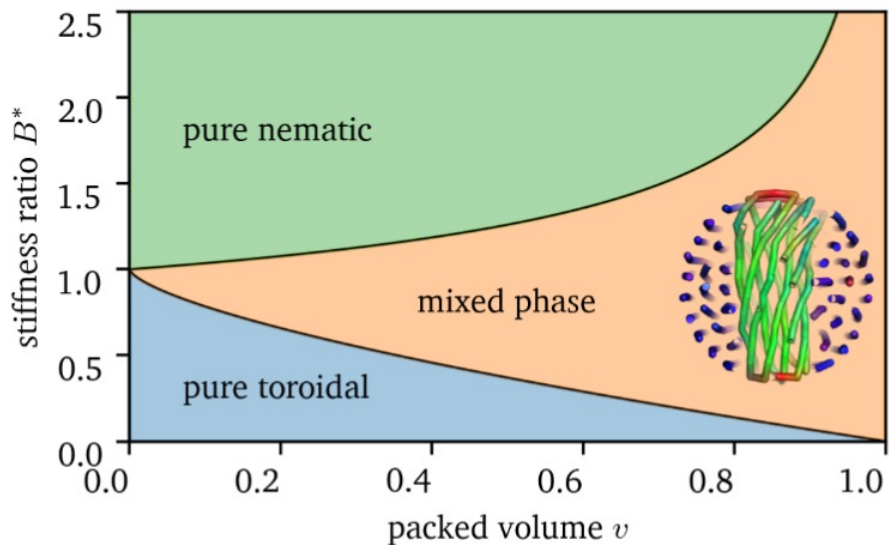
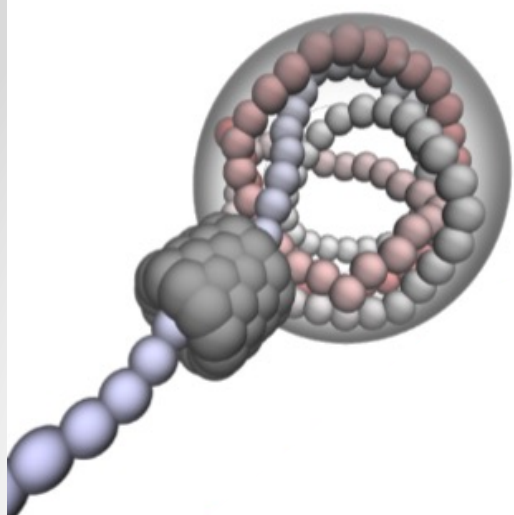
$$F(z, t) = Q(z) \sin(2\pi(z/\sigma - \omega t/\Delta t))$$

$$Q(z) = -4A \frac{(z - z_0)(z - z_1)}{(z_1 - z_0)^2}$$

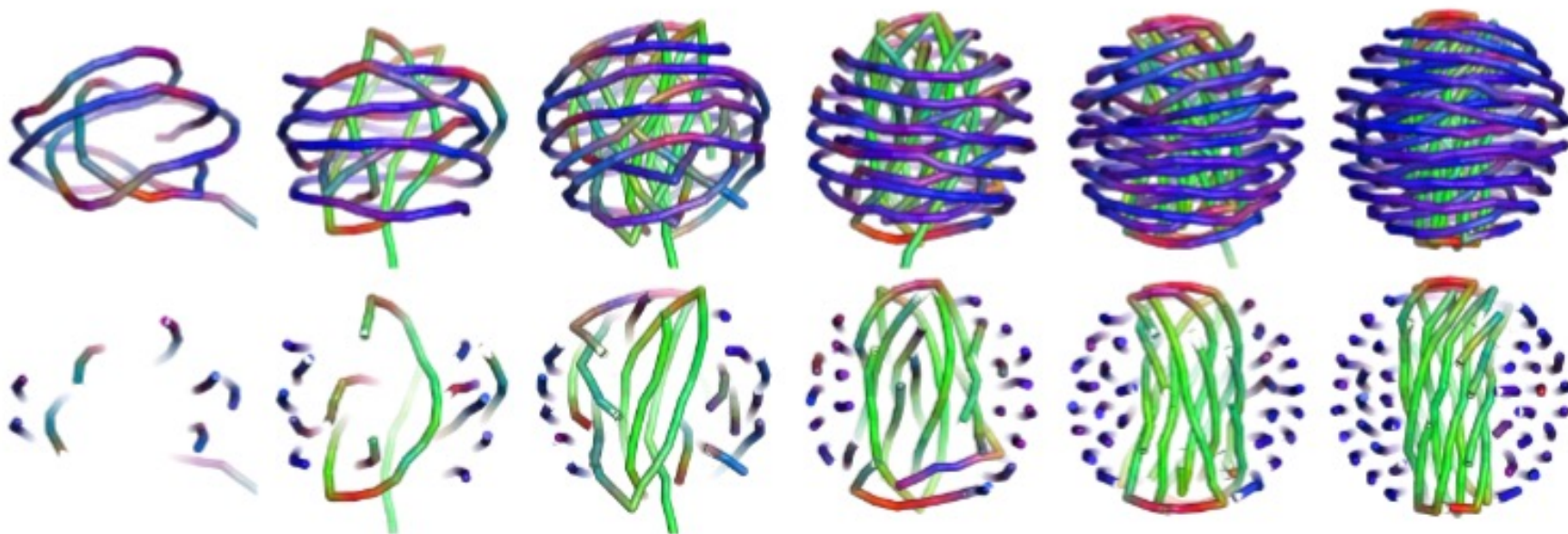


• **Sinusoidal packing motor with variable speed**

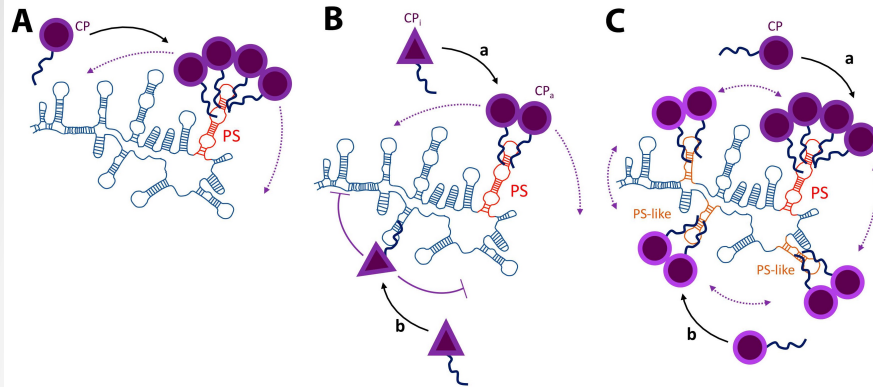




$$E = K \int_V |\mathbf{t}| \left( \hat{\mathbf{t}} \times (\nabla \times \hat{\mathbf{t}}) \right)^2 d\mathbf{r} + \varepsilon \oint_S |\mathbf{n} \cdot \mathbf{t}| dS$$



# Co-assembly of ssRNA viruses



## Co-assembly process:

- RNA partially hybridized
- ds segments + loops (PS)
- attractive interaction PS+CP
- icosahedral capsid sites (CS)



## Questions

- Osmotic pressure of packed RNA genome
- Effect of genome topology?
- Effect of capsid size?

## Hamiltonian Path Hypothesis:

self-assembly most efficient (min free energy)  
when adjacent PS bind adjacent CS

→ RNA: path visiting each vertex exactly once / forming order-1 bridges

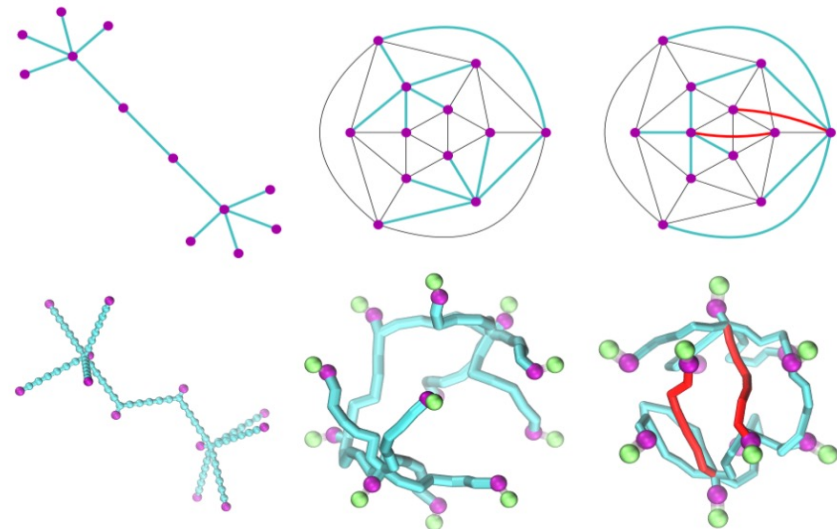
# Model

**CAPSID:** Spherical shell with 12 icosahedrally-distributed CS

## GENOME:

- Elastic chain with 12 regularly-spaced PS
- Encoded as graphs with adjacent PS as edges (trees)
- 551 unique trees
- 434 trees HPH-consistent

topology	linear	bolus	Wiener
MLD	11	5	5
$W$	286	194	176
$n_\theta$	10	22	18



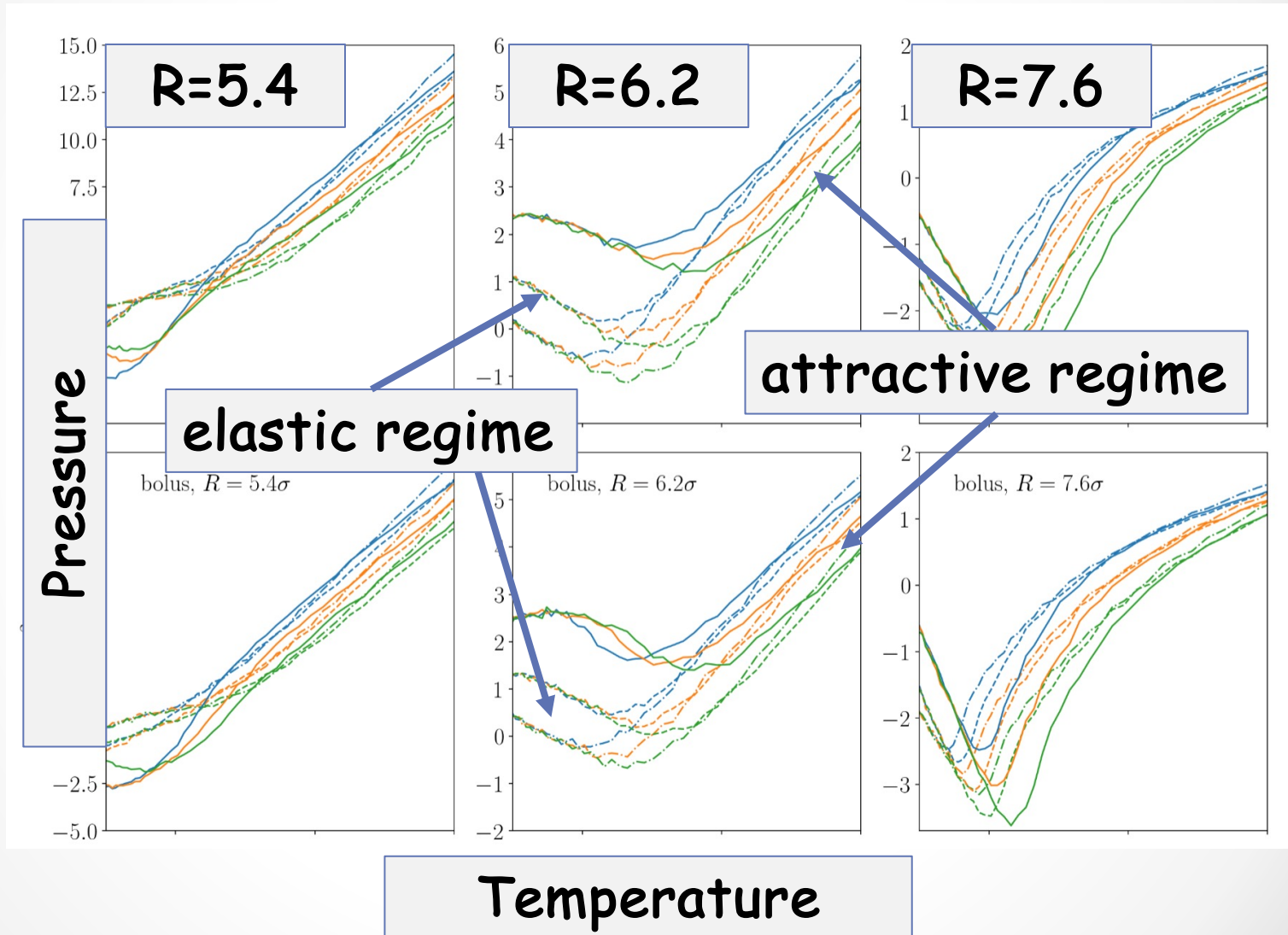
## INTERACTIONS:

- chain segments: harmonic  $V_b = 16 \Gamma (r - 1)^2$
- chain-capsid: repulsive WCA
- PS-CS: attractive (12-6 LJ)  $V_\theta = \Gamma (1 + \cos \theta)$

# Genome packing in ssRNA viruses

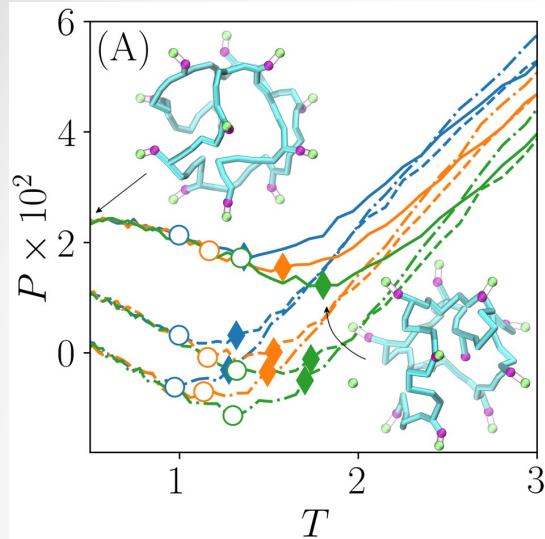
linear:  $n_{\theta}^{\min}$

bolus:  $n_{\theta}^{\max}$



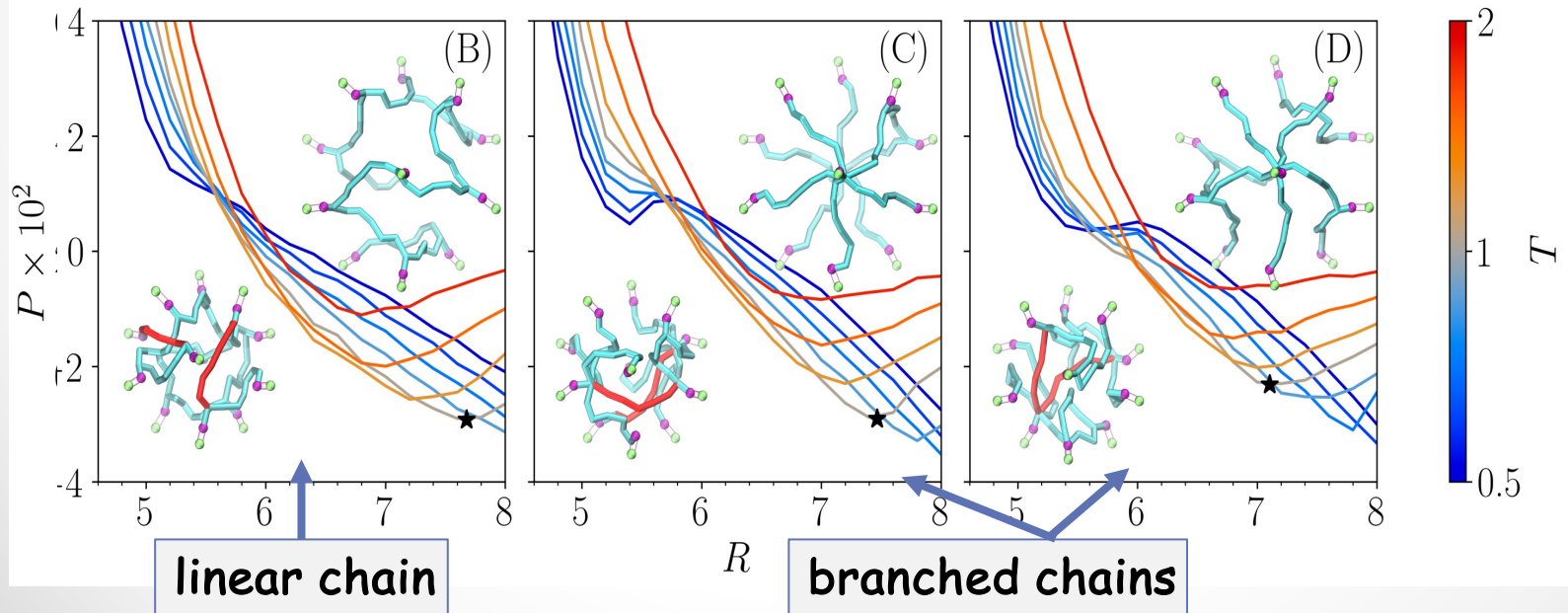


# Results: pressure in the capsid



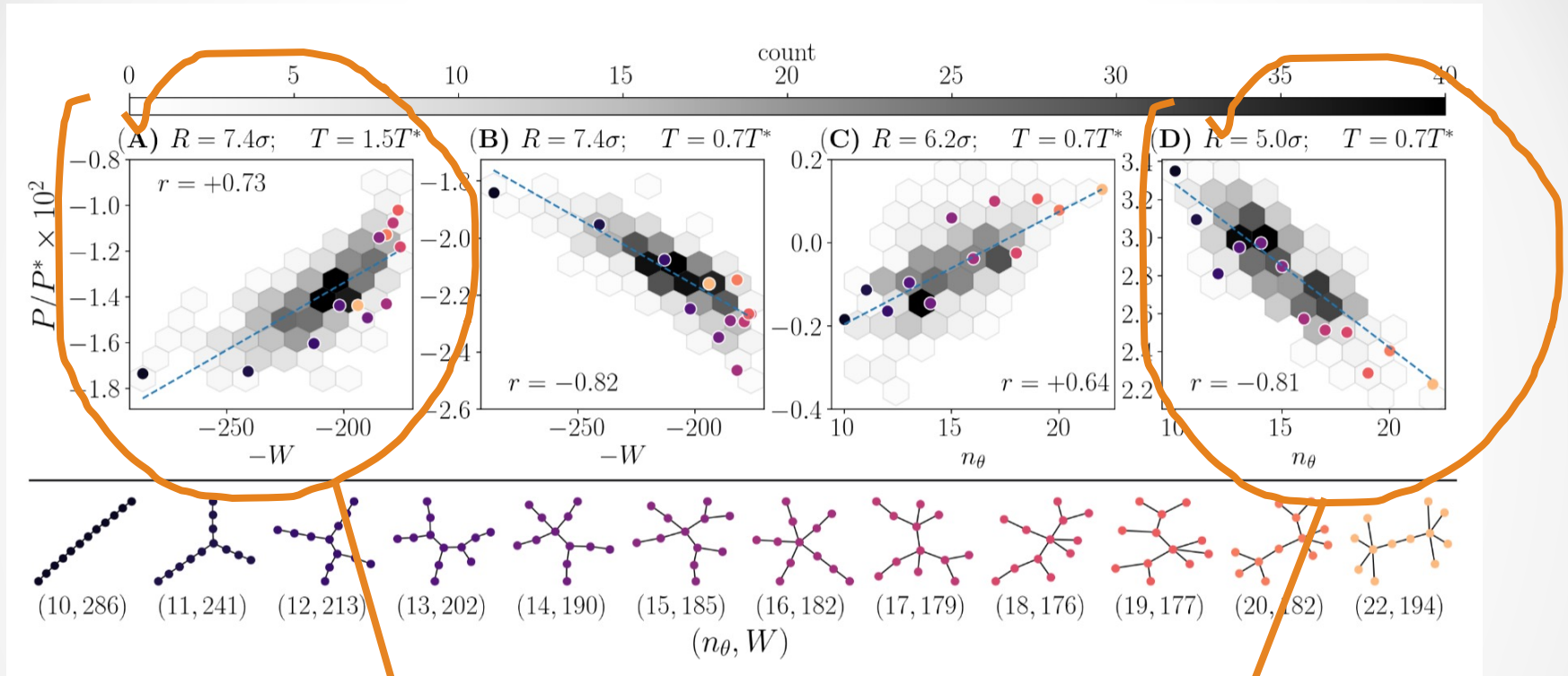
Pressure vs temperature:  
linear chain,  $R=6.2$

- blue/orange/green:  $\varepsilon_{BC}=12/14/16$
- solid/dashed/dot-dashed:  $\Gamma=1/2/4$
- **solid symbols:** 10/12 bonds formed (83%)
- **empty symbols:** 97% bonds formed





# Genome packing in ssRNA viruses

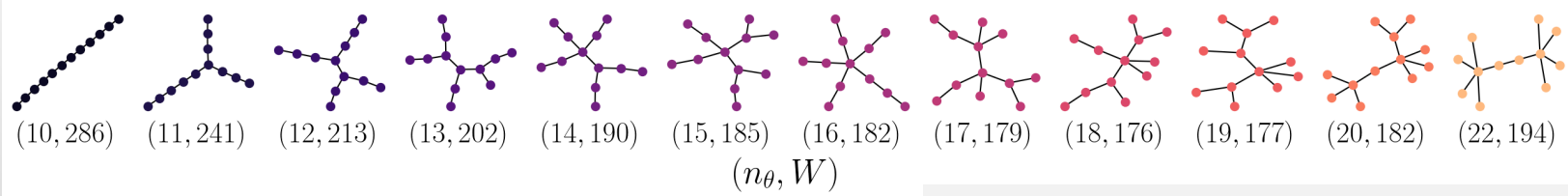
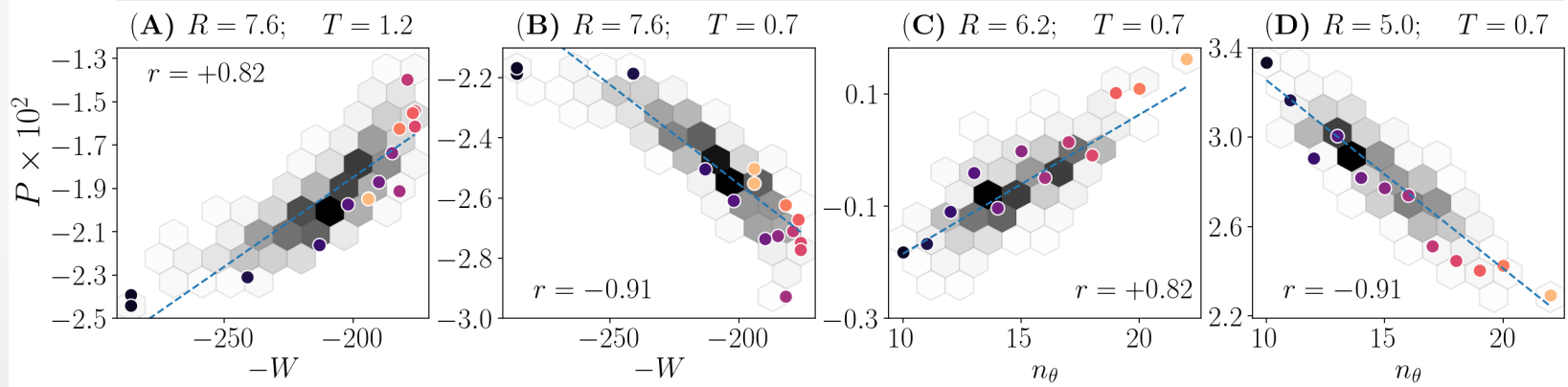
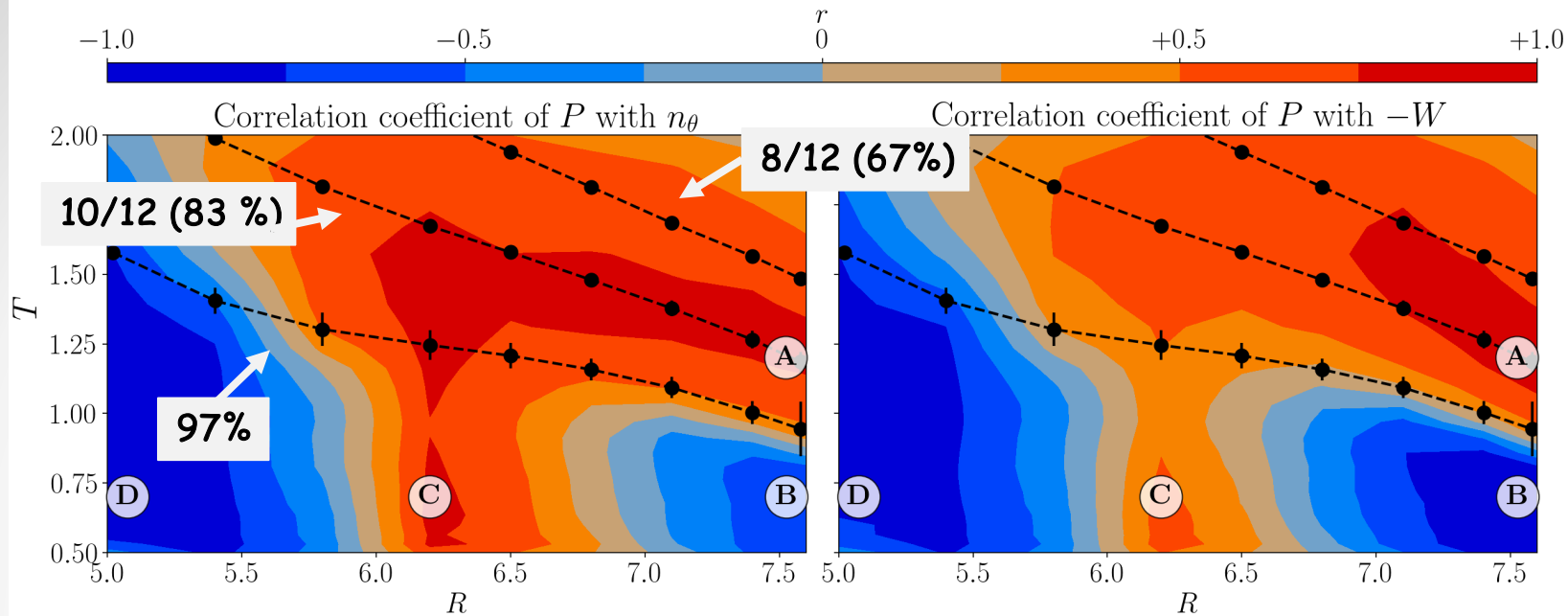


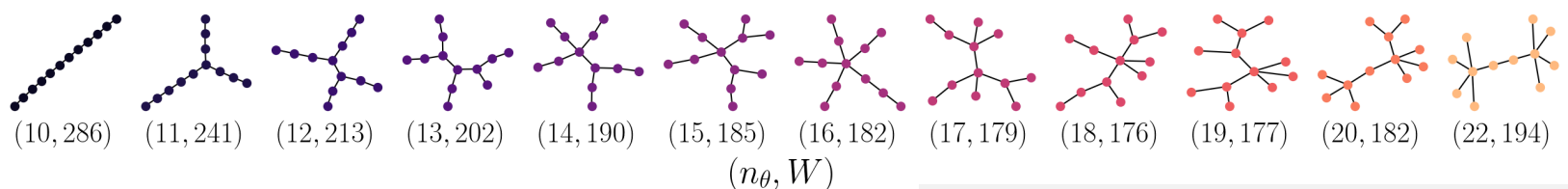
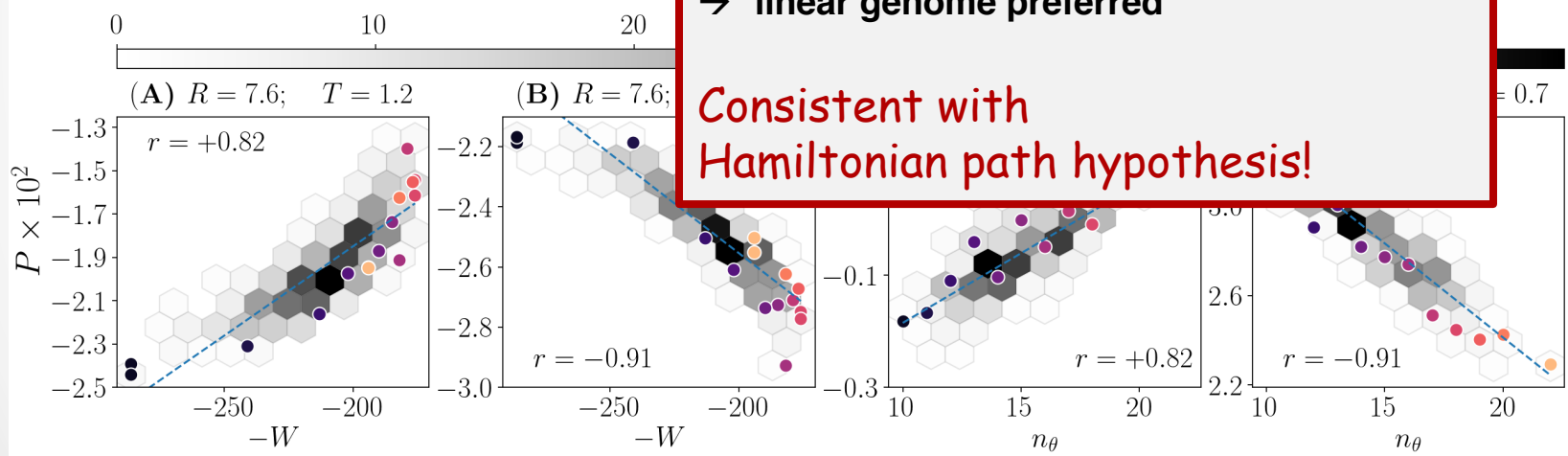
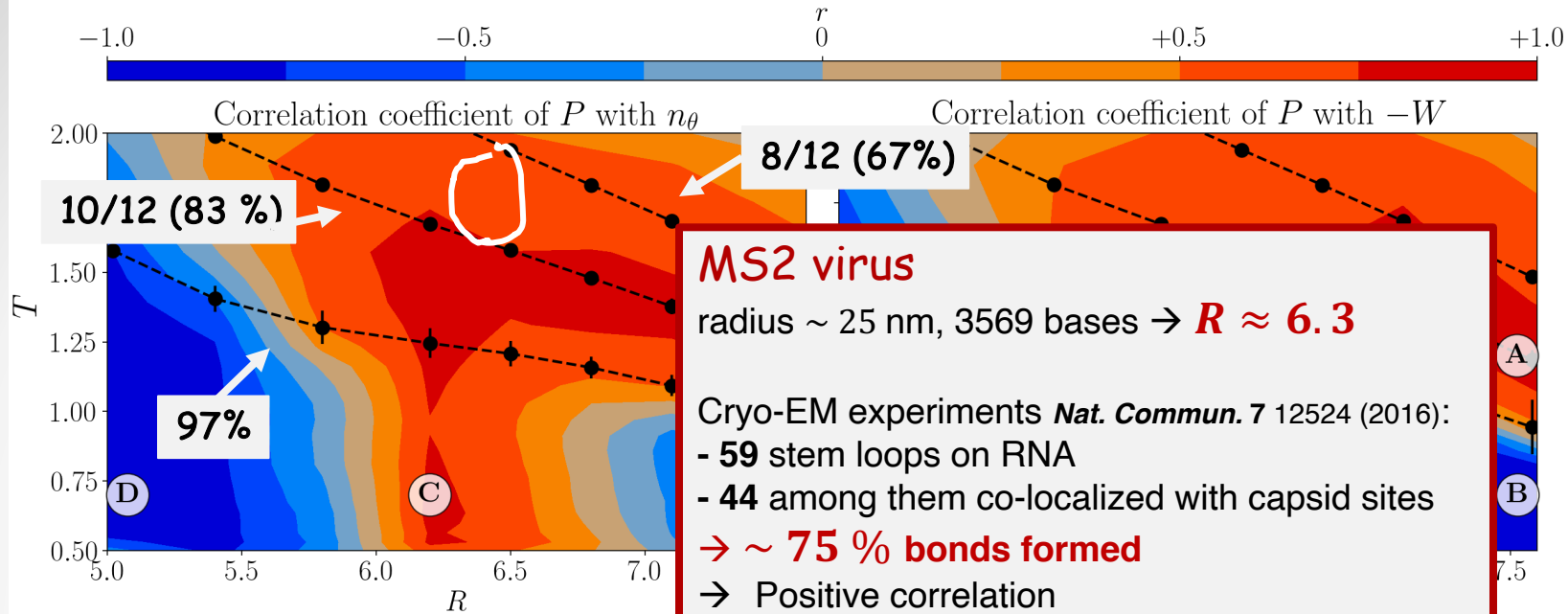
**LARGE CAPSIDS**

less branched genome  
more stable

**SMALL CAPSIDS**

highly-branched genome  
more stable





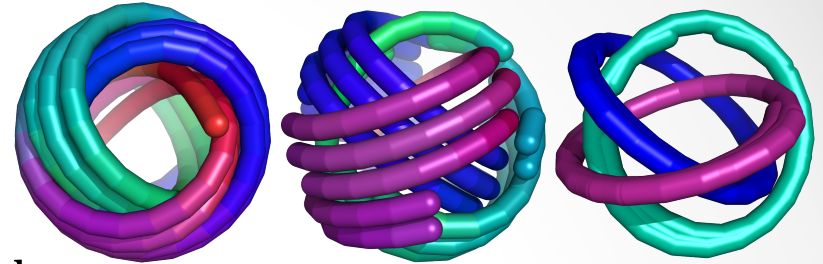
# Conclusions

Packing of elastic chains:

Multidomain arrangements

Spool-like solutions optimal

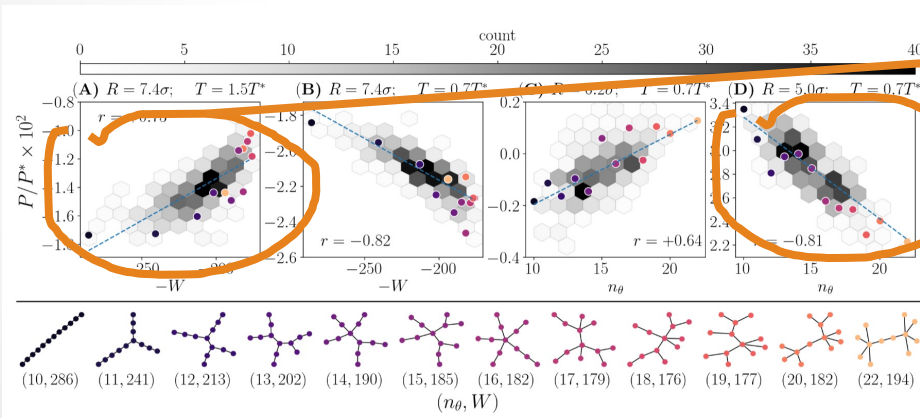
at low, **linked motifs** at larger chain lengths



Curk, Farrell, Dobnikar, Podgornik, *PRL* (2019)

Stability of co-assembled capsids with specific interactions:

Complex dependence on packing density and “temperature”



LARGE CAPSIDS  
linear genome

SMALL CAPSIDS  
branched genome

+ MORE COMPLICATED...

Farrell, Dobnikar, Podgornik, *PRR* (2023)

**Hamiltonian path hypothesis** seems justified for some viruses (not all)

**Two minima** (or multiple in more complex models) suggest coexistence of RNA conformations (*e.g.* **multipartite viruses...**).

Thank you.