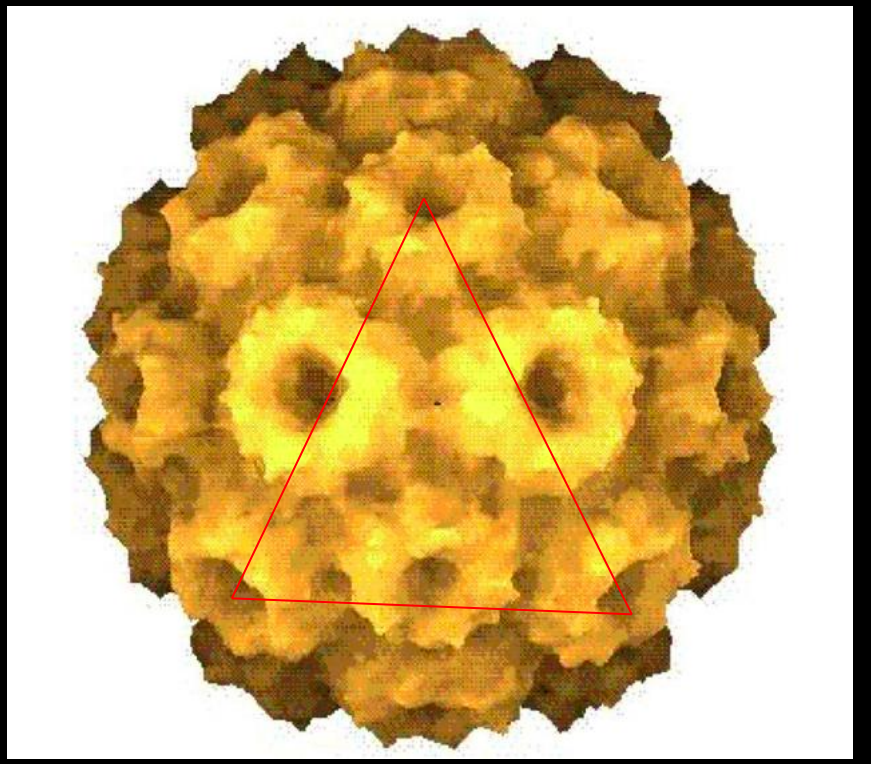
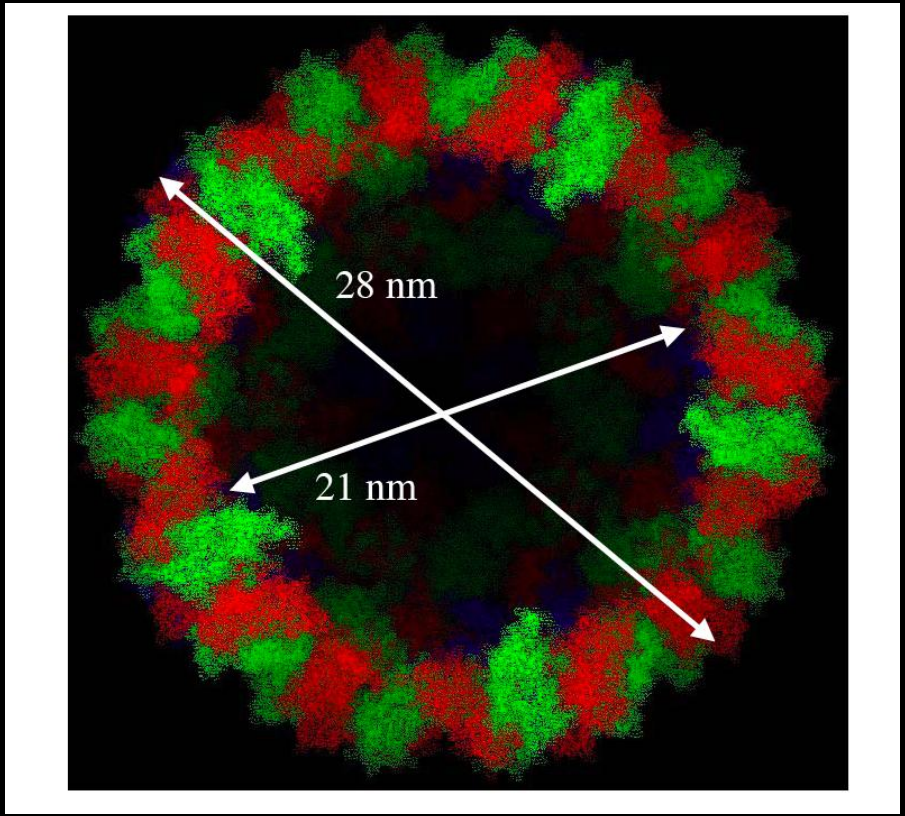
II) Fundamental Interactions.

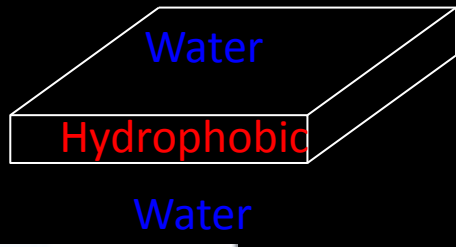
Cowpea Chlorotic Mottle Virus: CCMV



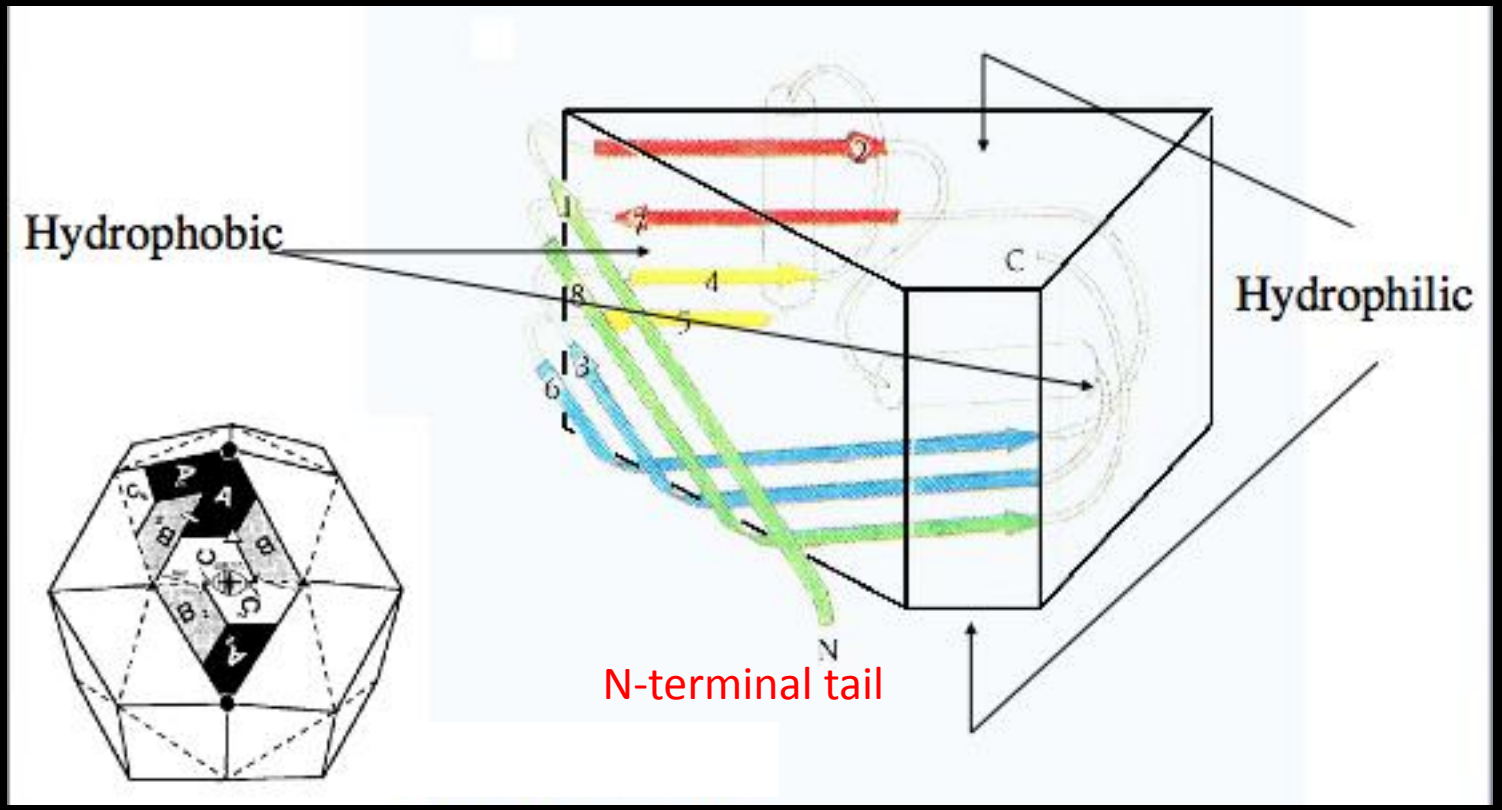


- Attractive Interactions

Layer capsid proteins



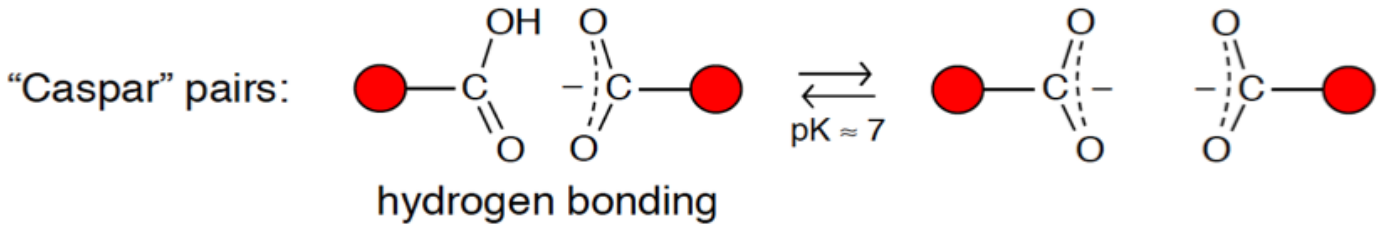
A) Capsid Proteins: Amphiphilic



- Layer: *Spontaneous Curvature*
- Expect reversible, thermodynamic assembly

B)

• pH sensitive “Caspar-pairs”.

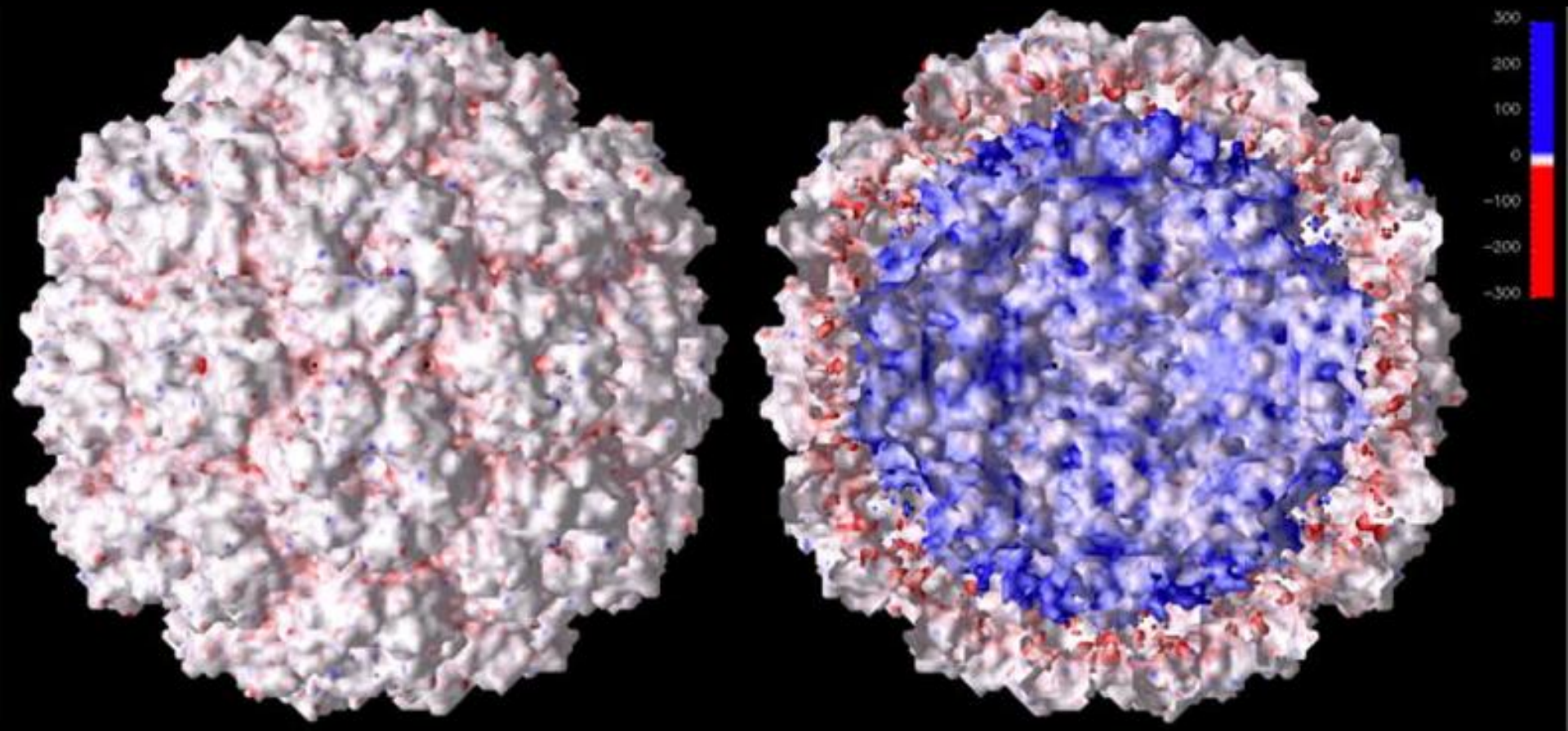


Kegel & van der Schoot,
Biophys J., 2006

- Strength attractive interactions increases with acidity.

Repulsive

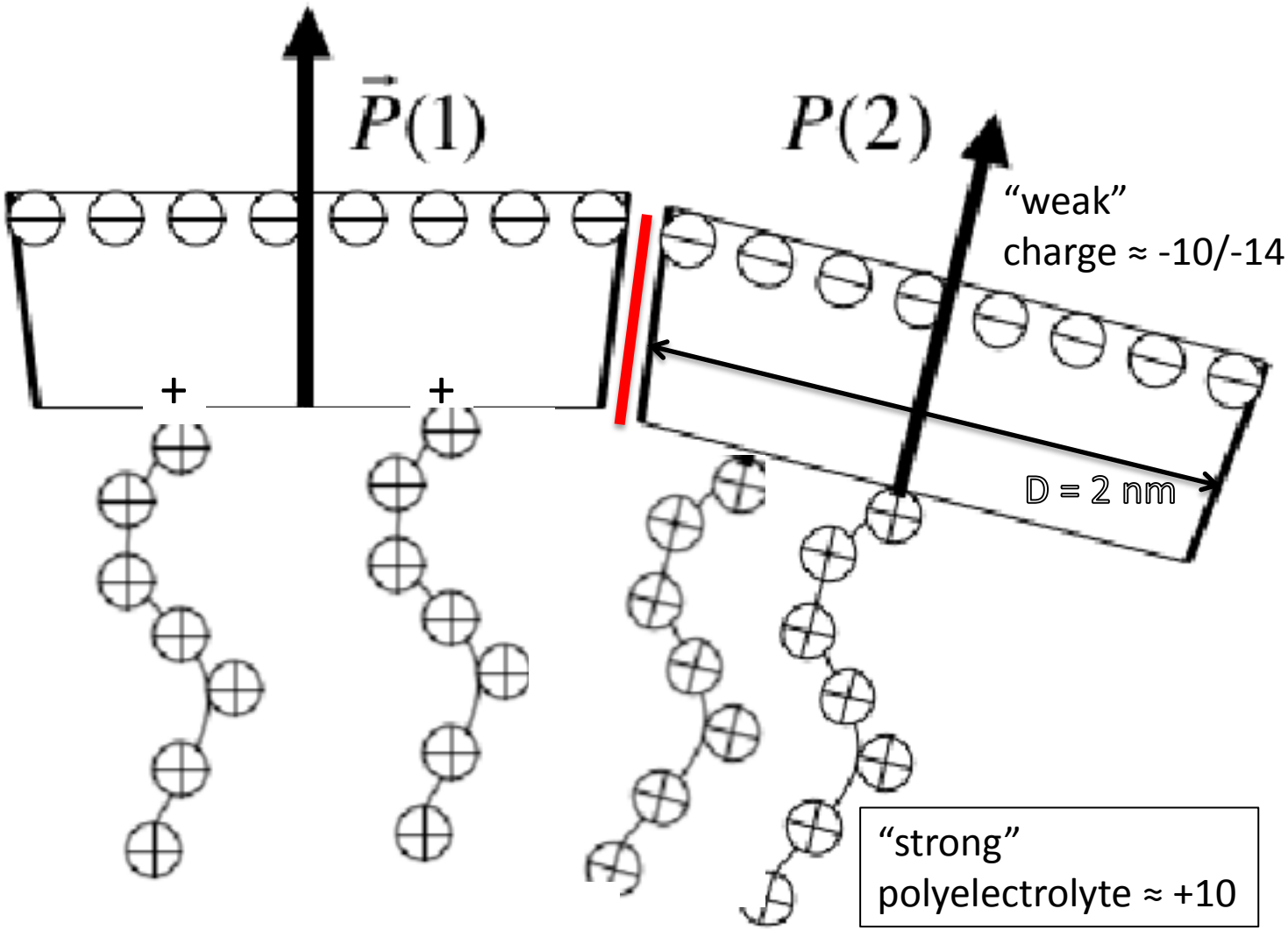
C) Electrostatic Interactions



- Water-accessible equipotential surfaces. Blue positive; Red negative.
- *Inside-Outside Voltage Difference*

(McCammon et al.)

Electrical Charges CCMV Dimers

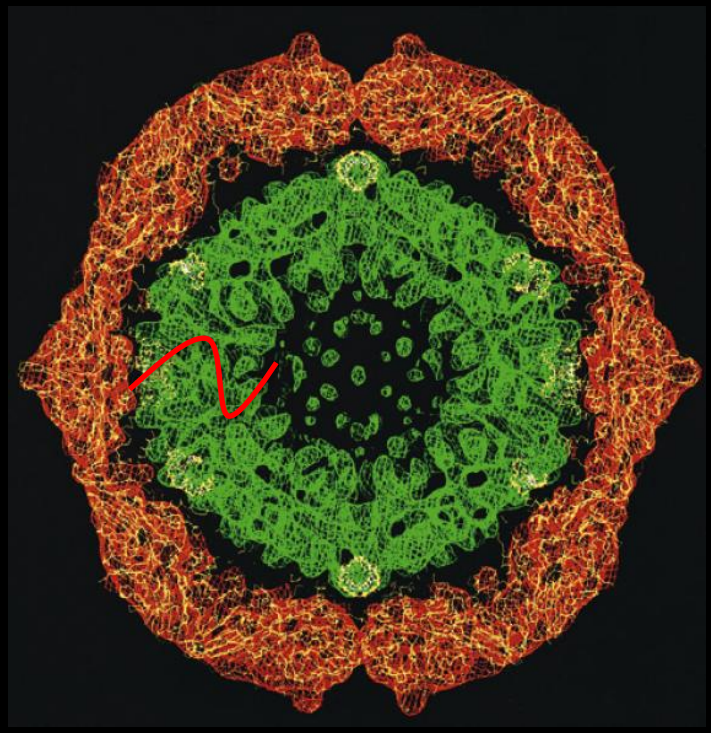


Some “just so” questions about CCMV electrostatics

RNA has a total negative charge ≈ -3,000

Positive tail charge ≈ 90 x 20 = + 1,800

- If tail/RNA neutralization promotes viral assembly, *why neutralize only a fraction of RNA charge* ? .



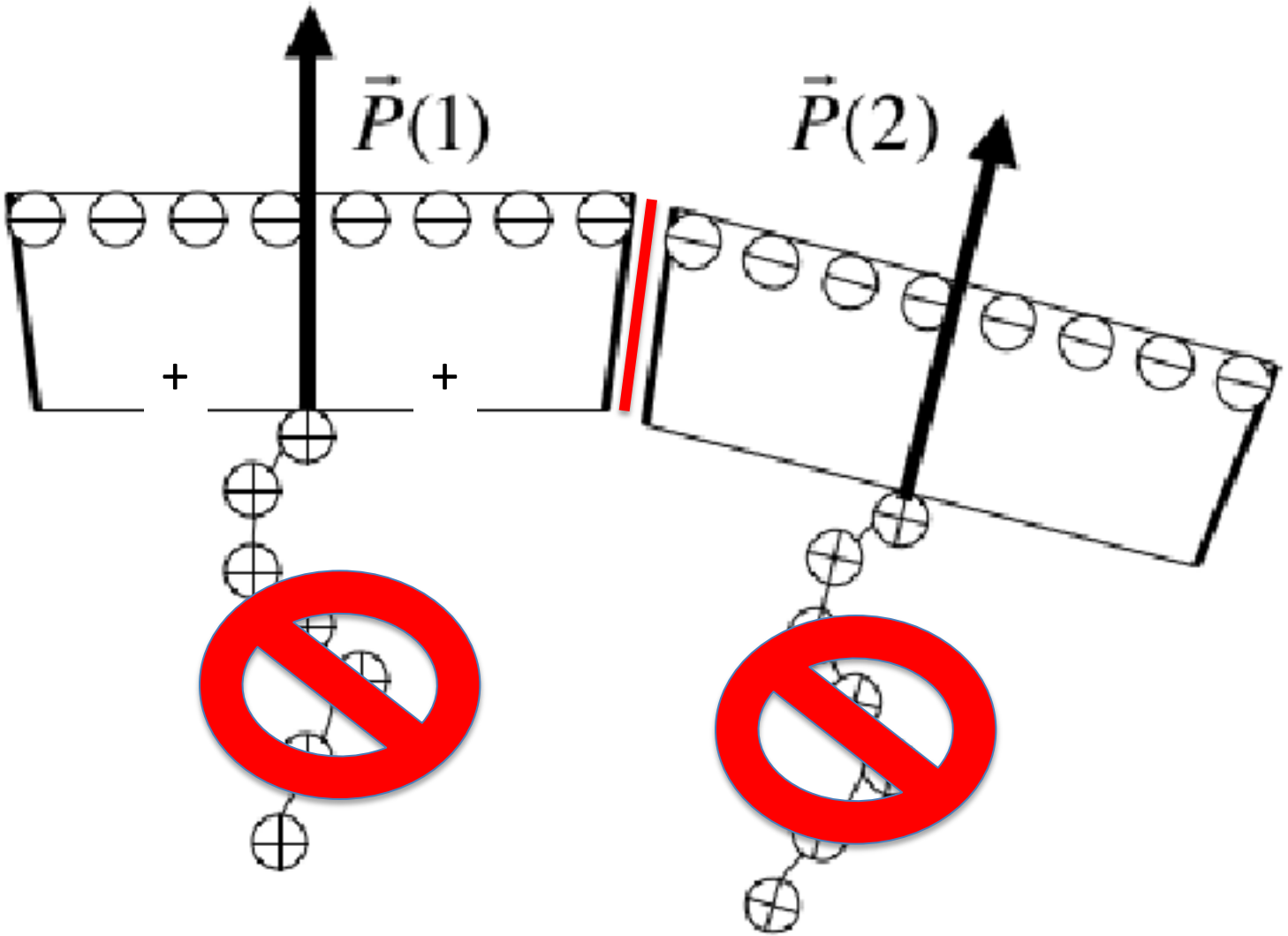
Outer layer charge ≈ - 28 x 90 = -2,520

2) *What’s the role of the large negative protein charge?*

- Prevents aggregation of viruses.
- Prevents RNA from sticking to capsids.

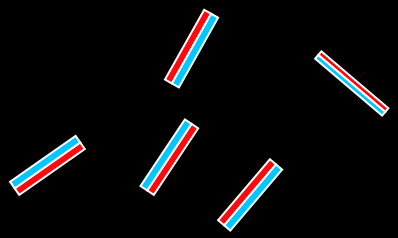
III) Self-Assembly Empty Capsids

Electrostatic Repulsion vs. Hydrophobic Attraction

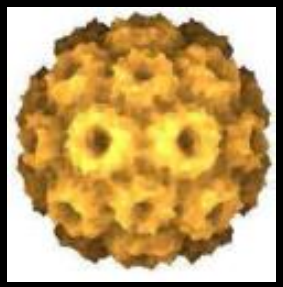


- Treat viral assembly as a *chemical reaction*:

90 Free CP Dimers ("subunits")



Assembled T=3 capsid



Thermal Equilibrium

$$dG = \sum_{components\ i} m_i dN_i = 0$$

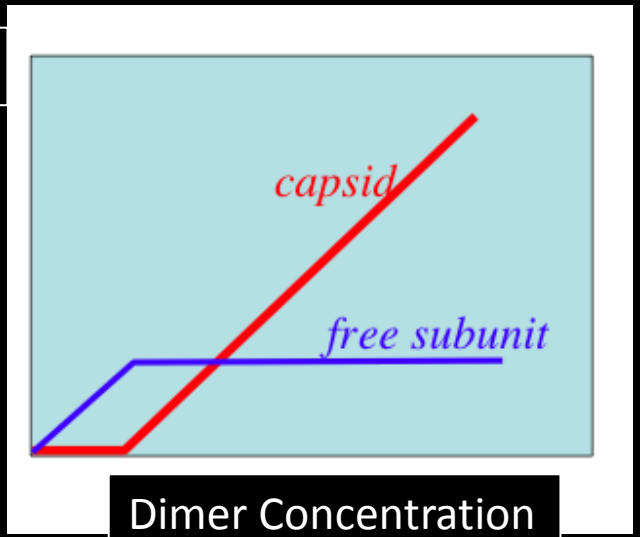
"Law of Mass Action"

$$K_{capsid} = \frac{[T=3\ capsid]}{[CP\ dimers]^{90}}$$

$$K_{capsid} \propto \exp(90\Delta G / k_B T)$$

- ΔG = assembly energy/dimer

Concentrations



"Signature" of Thermodynamic Self-Assembly

Empty capsid assembly experiments

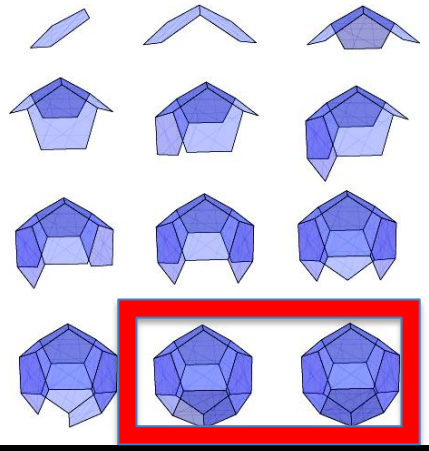
* Acidic environment (low pH)

$$dG = 0$$

$$K_{capsid} = \frac{[T=3\ capsid]}{[CP\ dimers]^{90}}$$

- $\Delta G \approx 30\ k_B T$/dimer.

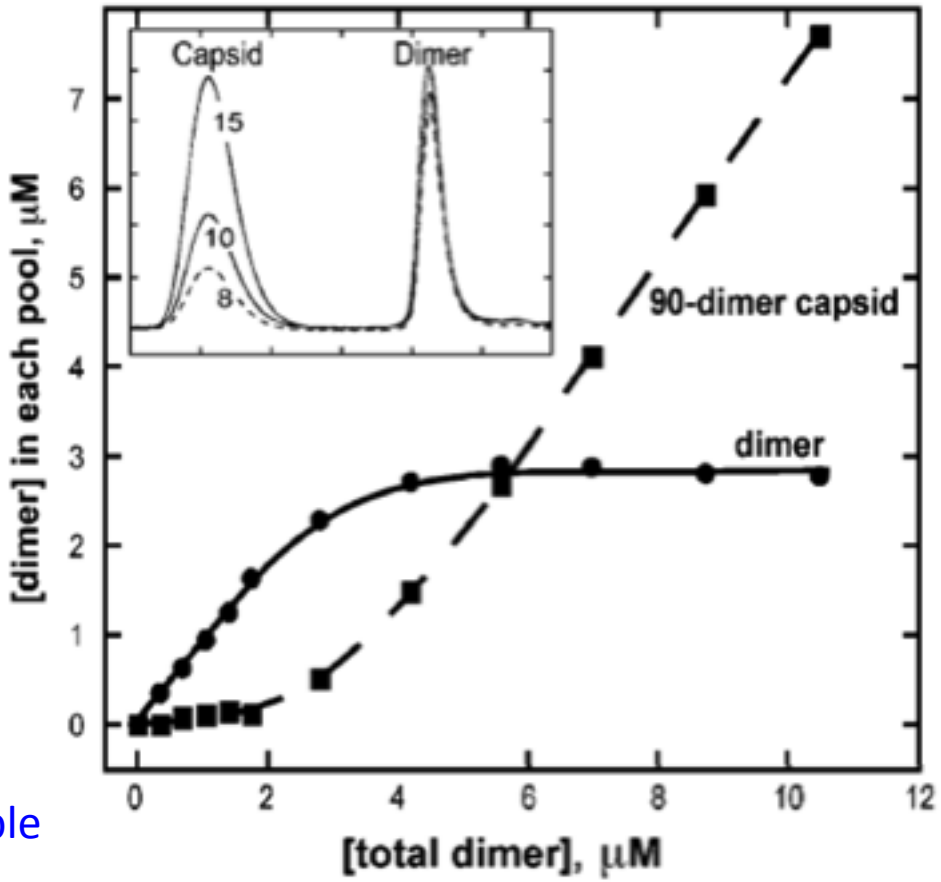
- Capsid assembly is *irreversible!?*



reversible

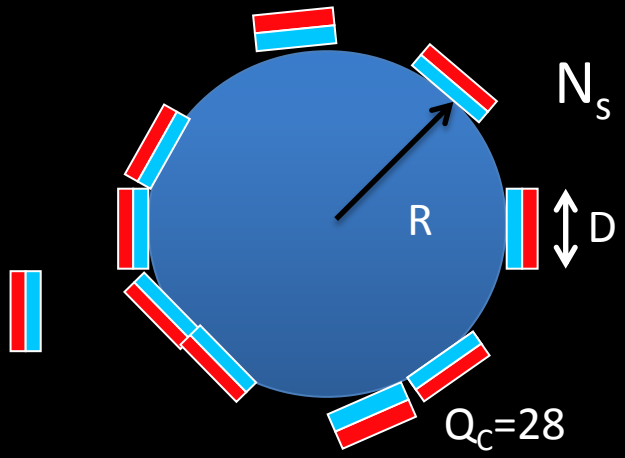
irreversible

Chromatography



Zlotnick (2005) Nano Letters **5**, 765-70

Capsid Van der Waals/Landau Free Energy



N_s adsorbed proteins

$\rho_s = [N_s / R^2]$ area density "order parameter"

Entropic Free Energy 2D ideal solution

$$F_S(N_s,R)/N_s \gg k_B T \ln r_s D^2 - \varepsilon + v_s r_s + w_s r_s^2 + \ldots$$

ε: Adsorption energy proteins on sphere.

v_s: Second "virial coefficient"

w_s : Third virial coefficient ≈ $k_B T\ D^4$

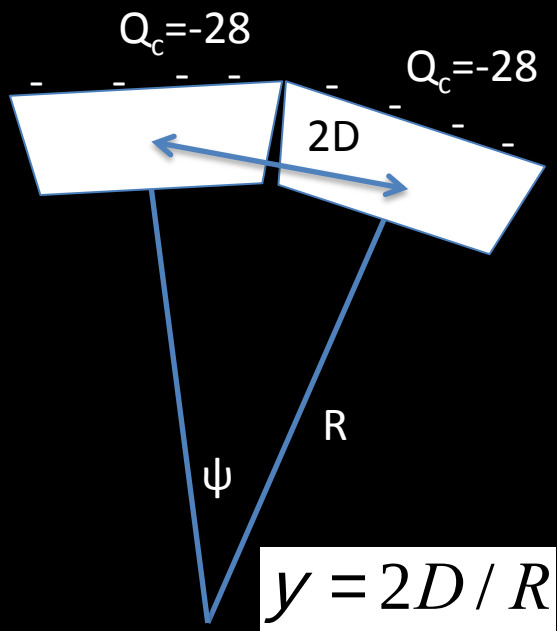
Thermal equilibrium:

$$\frac{\partial F}{\partial N_s} = \mu$$ Chemical potential proteins

Second Virial Coefficient

- Electrostatics vs Hydrophobicity

$v_S = v_{DH} - J$



Capsid Proteins

$y = 2D / R$

"Bjerrum Length" ≈ nm

$$v_S(y) / k_B T = \frac{l_B Q_C^2}{k} - J_S(y)$$

"Debye Parameter" ≈ 1/nm

Angle-dependent hydrophobic attraction

Optimal Angle/ Radius

$y_c = 2D / R_c$

CCMV (pH=5): $Q_C = 20$

$v_{DH} / k_B T = 400\ nm^2$

$v_S(y = y_c) / k_B T \approx -70\ nm^2$ Measured for empty shells

$J_S(y = y_c) / k_B T \approx -470\ nm^2$ CP-CP Hydrophobic Attraction

Capsid Radius

Debye-Hückel Theory of Aqueous Electrostatics

Macro ion Charge Density (CPs/RNA)

$$\Delta\phi - \kappa^2\phi = -\frac{4\pi e\rho_M(\vec{r})}{\epsilon}$$

Dielectric Constant Water

Electrical Potential

Debye parameter

$$k^2 = 4\pi e^2[Salt]/\epsilon k_B T$$

Electrostatic Free Energy

Bjerrum length

$$\frac{e^2}{\epsilon l_B} = k_B T$$

Sheet of charges

$$F_{DH} = \int d^3r\left(\frac{1}{2}e\rho_M(\vec{r})\phi + \frac{\epsilon}{8\pi}\kappa^2\phi^2\right)$$

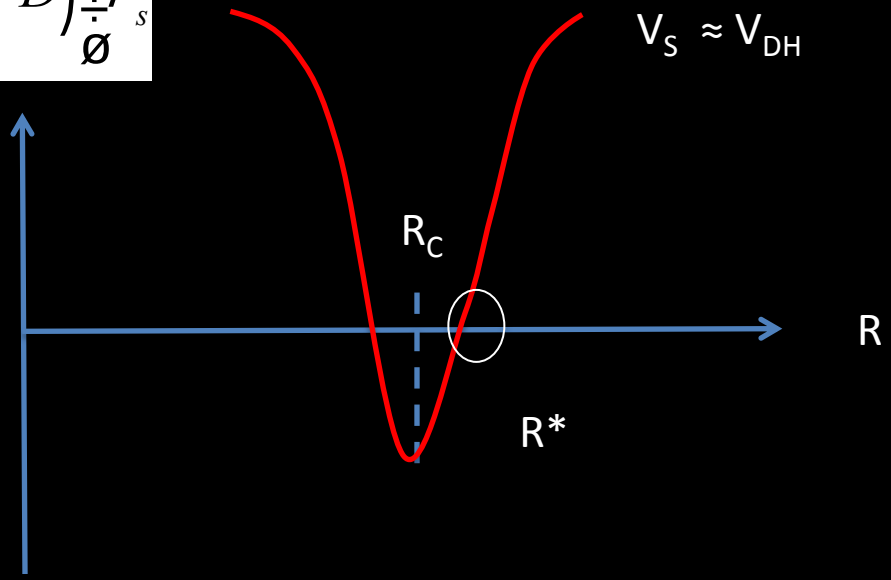
$$F_{DH}/k_B T = \left(\frac{l_B Q_C^2}{k}\right)\frac{N_S^2}{R^2}$$

Summary

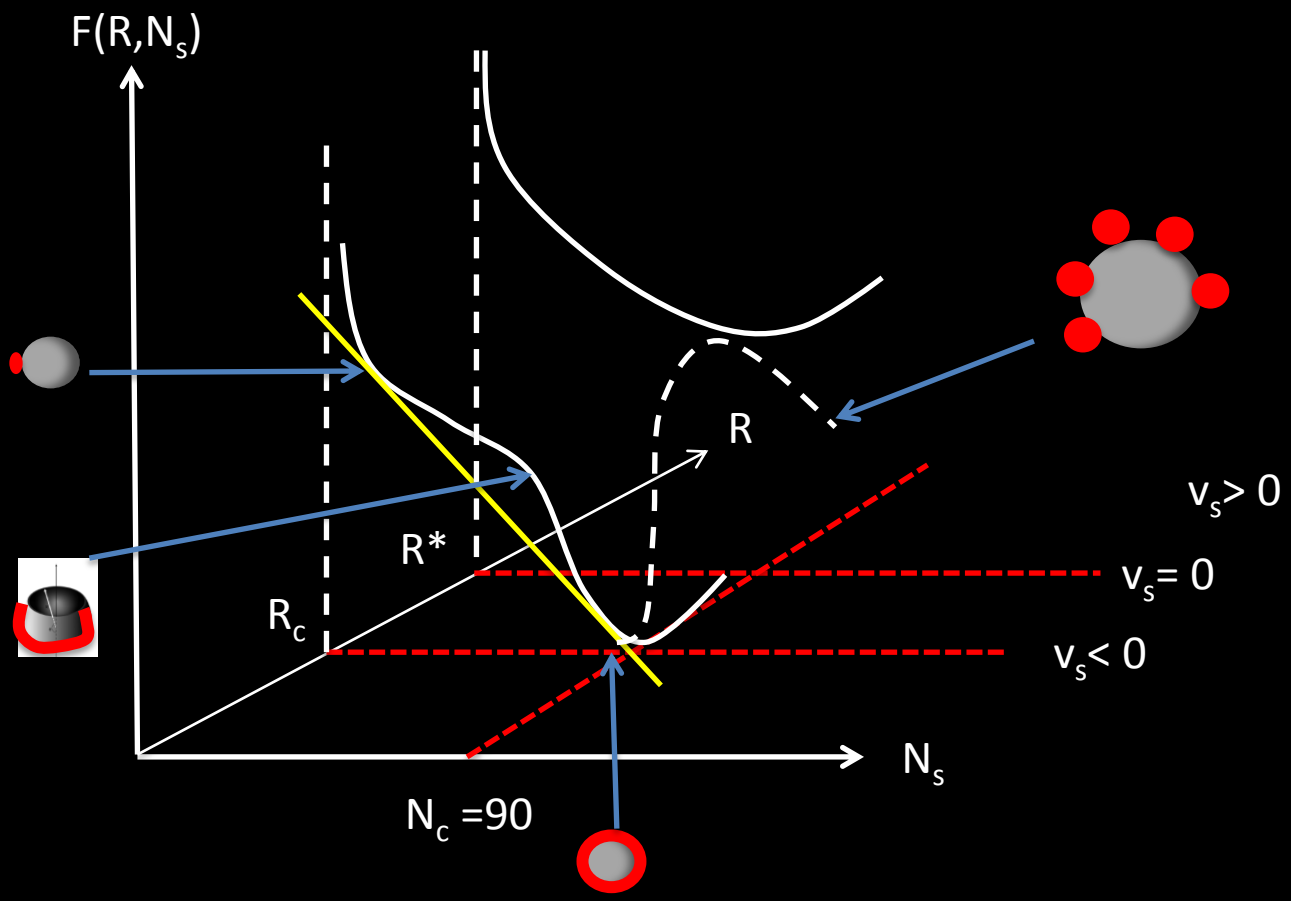
- *Delicate balance between large repulsive interactions and large attractive interactions*

- *Second virial coefficient depends on the sphere radius R.*

$$F_2 / N_s = k_B T \left(\frac{q_B Q_C^2}{k} - J_S\left(R_c / D\right) \right) r_s$$



Free energy “landscape” $F(R,N_s)$



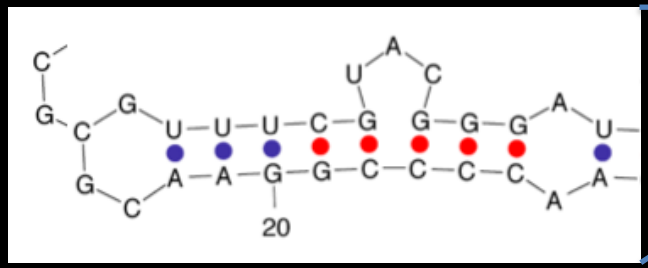
$$\frac{\P F}{\P N_s} = m$$

“Common-tangent construction”

Phase-coexistence: nearly closed shells and nearly bare spheres

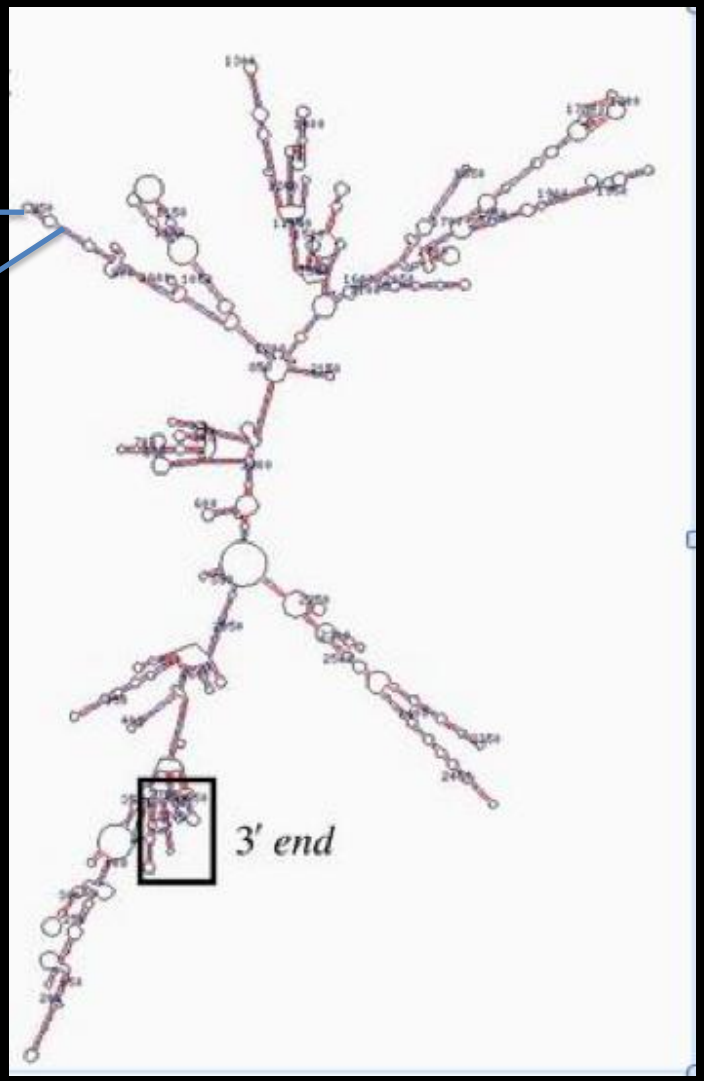
IV) Condensation RNA genome molecules

- Highly branched, highly charged “polyelectrolyte”



Paired stretches l ≈ 5 bp

≈ 100 nm



Neutron scattering R ≈ 11 nm
(no “condensing agents”)

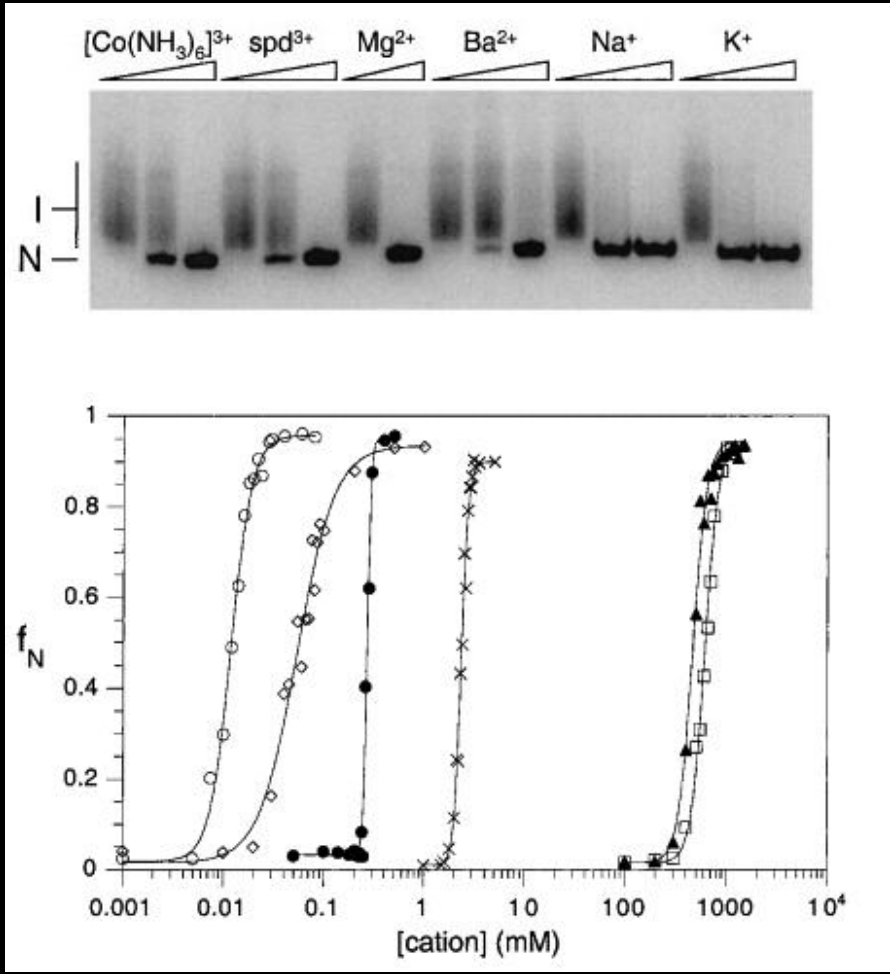
- Highly compactified

(Knobler et al.)

CCMV RNA 1

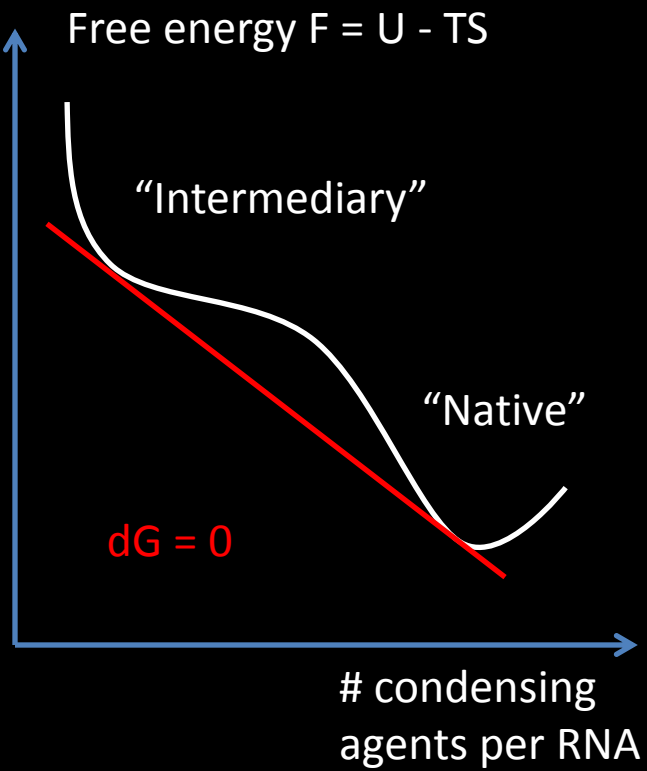
RNA Condensation

Condensing Agent



"Folded Fraction"

Condensing agent concentration (polyvalent counterions)



Gibbs Free Energy

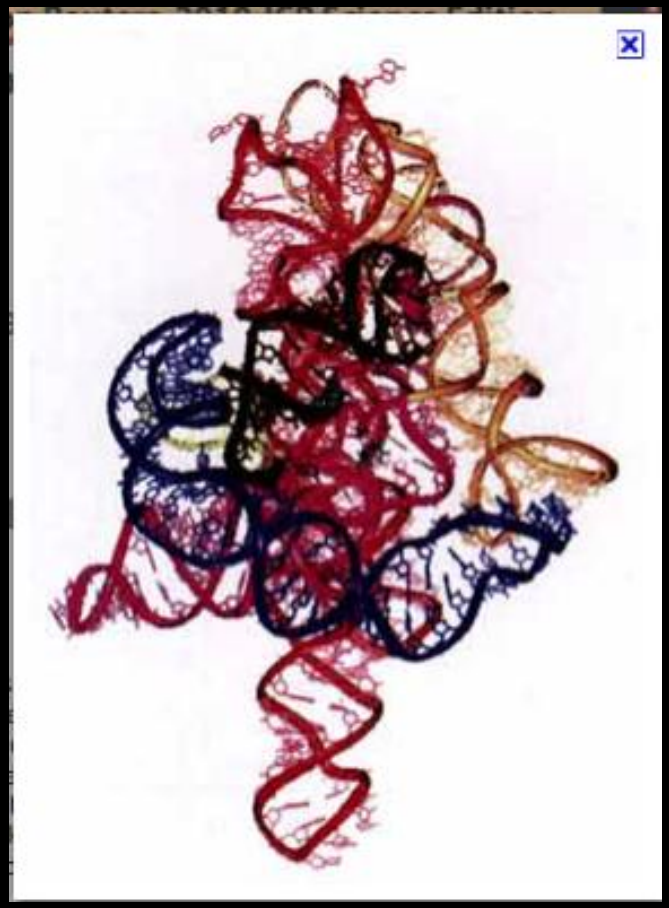
$G = F - \mu([\text{agent}])\,\theta$

Chemical potential condensing agent

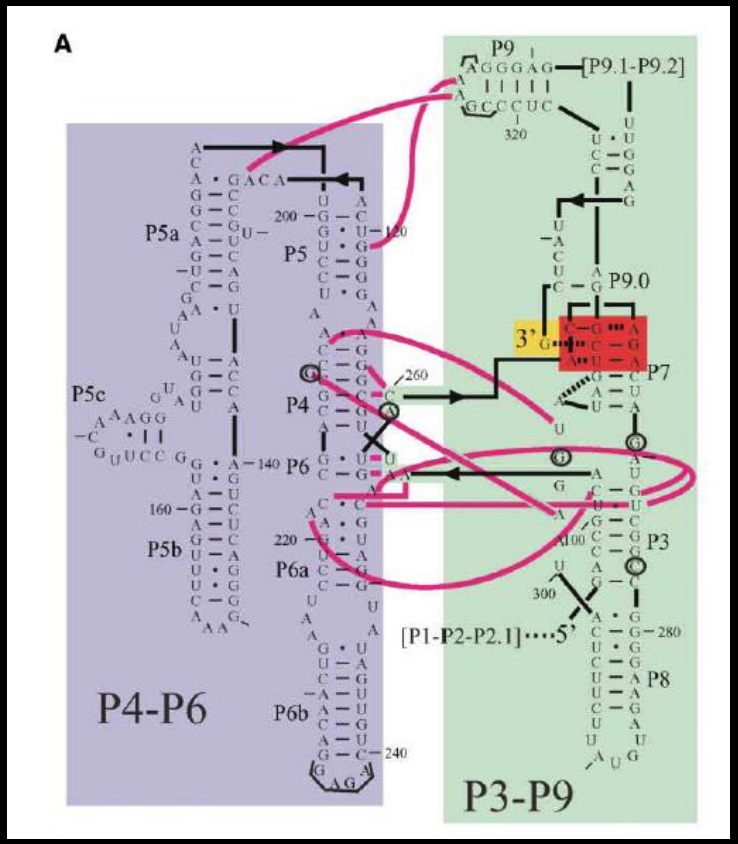
- Highly cooperative, first-order phase transition.

Koculi E, Lee NK, Thirumalai D, Woodson SA.
J Mol Biol. 2004, 341(1):27-36.

N state: folded

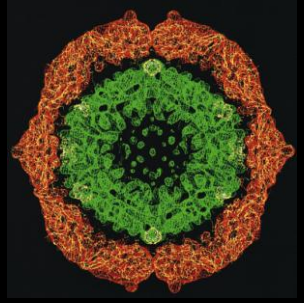


Tertiary contacts



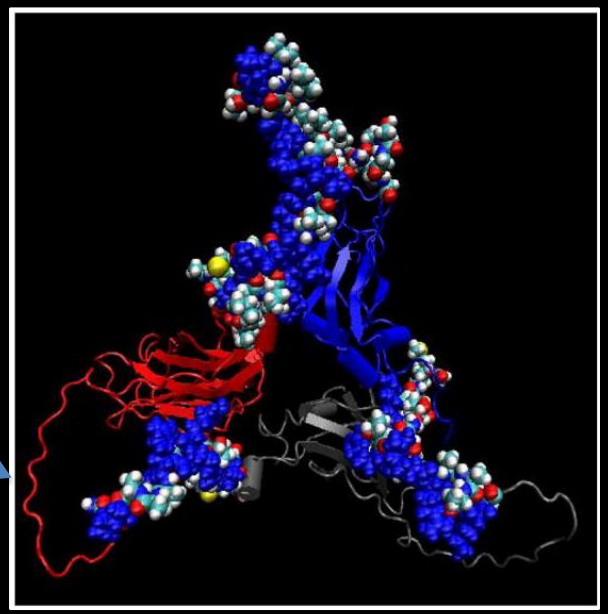
- Ribozyme (Tetrahymena) RNAse (Cech)

- RNA inside T=3 virus:
  Highly condensed



What are the condensing agents ?

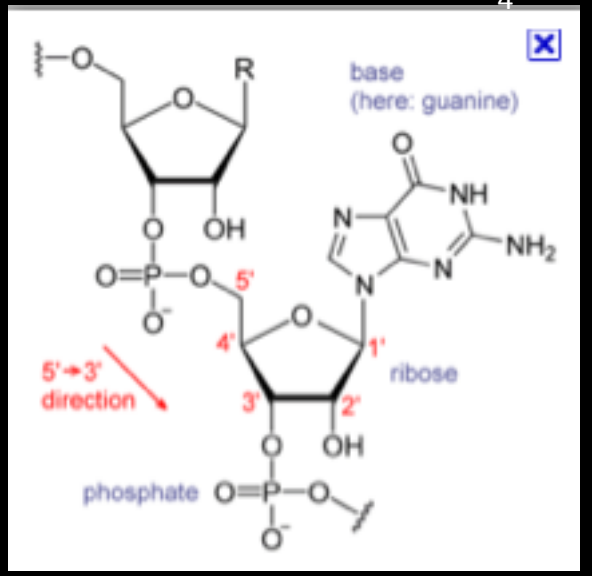
CCMV Capsid protein



Disordered N-Terminal Tail:
+ 10 charges

RNA Condensing Agent

ss RNA PO_4^-



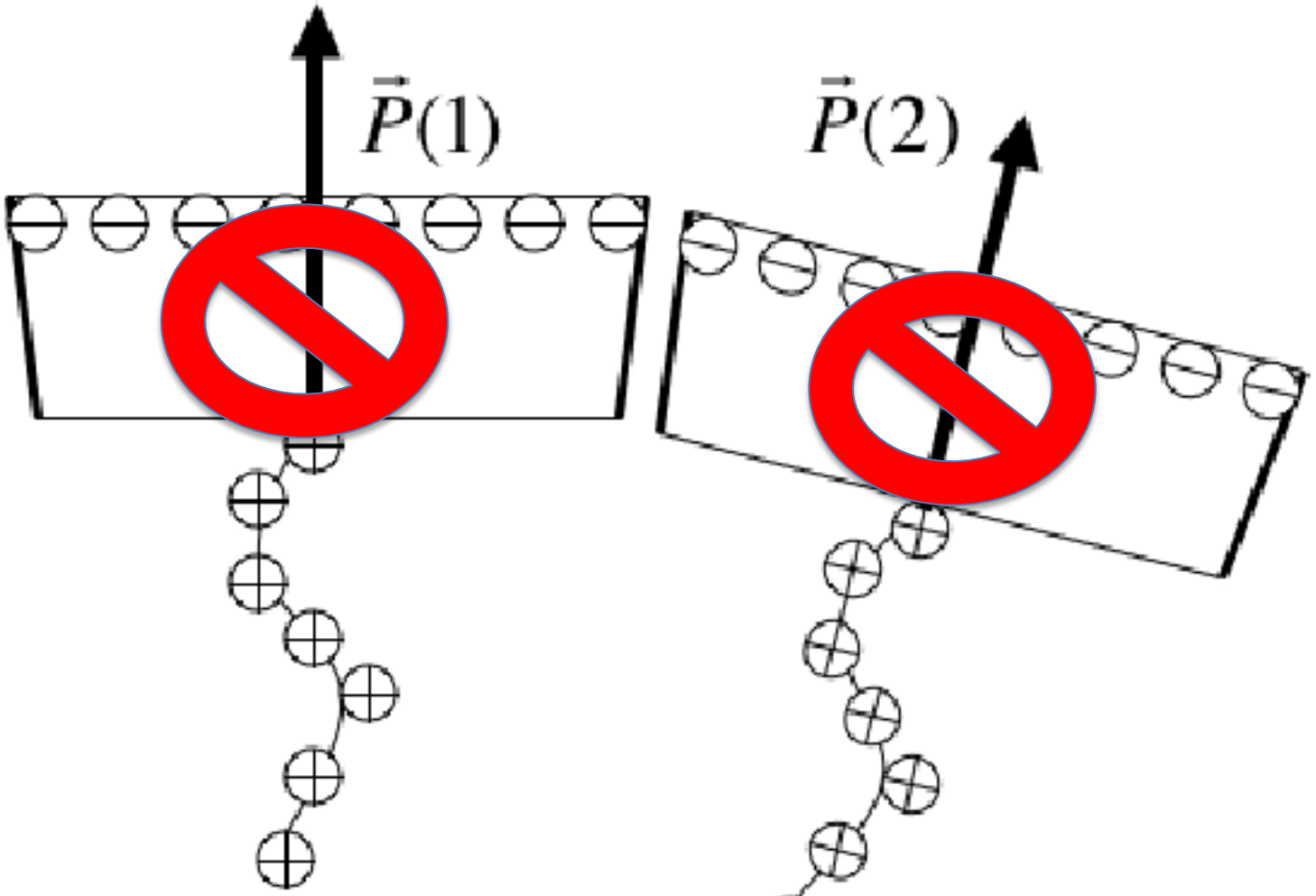
Numerical Simulation: ε (tail/RNA)≈ 10-15 $k_B T$

Zhang et al.

Biopolymers. 2004 November; 75(4): 325–337)

Remove Protein Cores

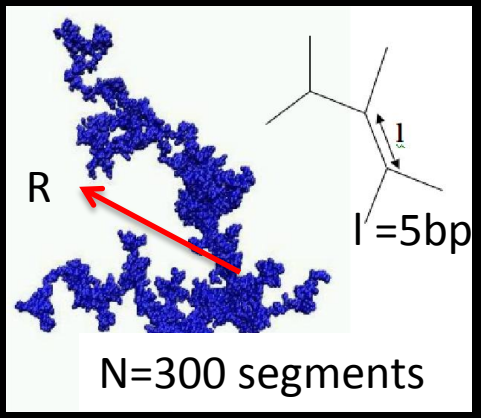
CCMV Dimers



Q_T=+20 tail charges/dimer

Condensation of CCMV RNA

Good Solvent: "fractal"
$R(N) \approx N^{1/2}$



- Flory-Landau mean-field theory

Entropic Elasticity U

$$F_F(R,N) = k_B T \frac{R^2}{R_0^2} + V(q)\frac{N^2}{R^3} + W\frac{N^3}{R^6} + ..$$

$$R_0(N) = l\, N^{1/4}$$

Radius gyration of an "ideal" Flory-Stockmayer branched polymer

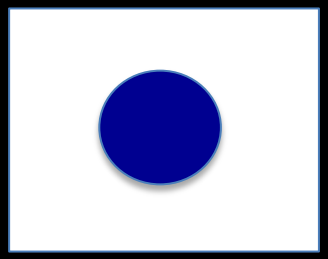
Linear polymers: much larger

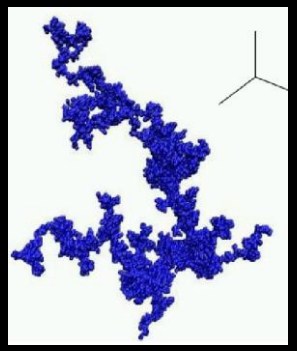
$$R_0(N) = l\, N^{1/2}$$

- V(θ): Second Virial Coefficient. θ: # tails / segment
- W: Third Virial Coefficient $\approx k_B T\ l^6$

Condensed Globule

Swollen Fractal





V

$$R(N)/l \approx \left(\frac{V}{W}\right)^{1/3} N^{1/3}$$

V=0

"Theta Solvent"

$$R(N) \approx \left(V\, l^2\right)^{1/3} N^{1/2}$$

CCMV RNA genome free in solution

R ≈ 11 nm

- l = 0.5 nm
- $V(\theta=0)/\ k_B\ T$ = 1-10 nm^3

* No phase transition

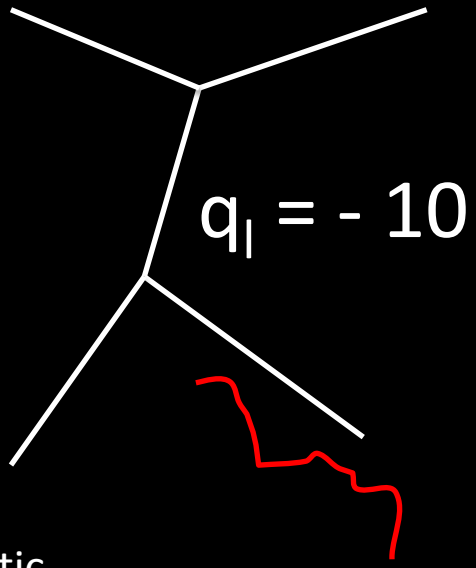
Second Virial Coefficient

Segment charge

Bjerrum Length

Tail fraction

$$V(q)/k_BT = \frac{3l_B}{k^2}\left(-\bar{q}_l + Q_T q\right)^2 - \bar{V}_{RNA}$$



Maximum concentration

$$q_{\max} = q_l / Q_T \approx 0.5$$

Neutralization

Debye parameter

Polyvalent Counterion Charge

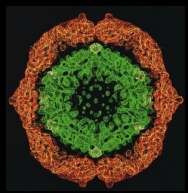
Non-electrostatic

$$V(q=0)/k_BT \approx 100\ nm^3 - \bar{V}_{RNA} \approx 1 - 10\ nm^3$$

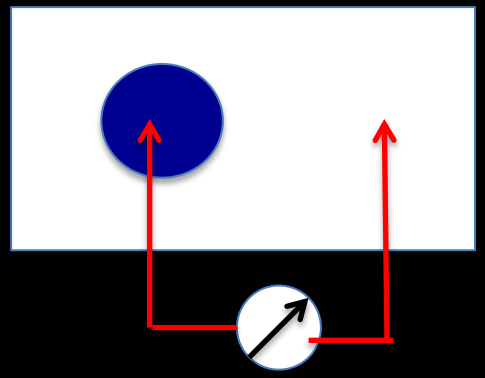
(free RNA in solution)

$$V(q=q_l/Q_T)/k_BT = -\bar{V}_{RNA} \approx -100\ nm^3$$

RNA/tail association:
"unveils" strong RNA self-attraction



RNA Globule



Voltmeter

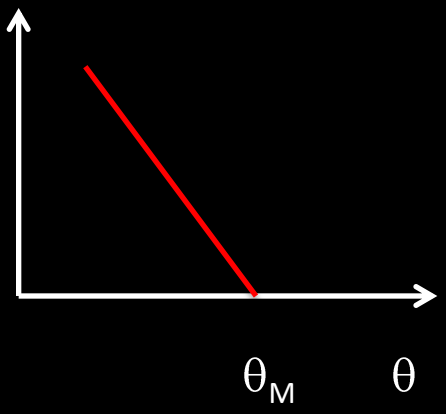
ΔV

$$e\mathrm{D}V_D(q)/k_BT = \frac{l_B\left(-q_l + Q_T q\right)}{k^2 R^3}$$

"Donnan Potential"

$V(q)$

15-20 milli Volt



Charge neutral

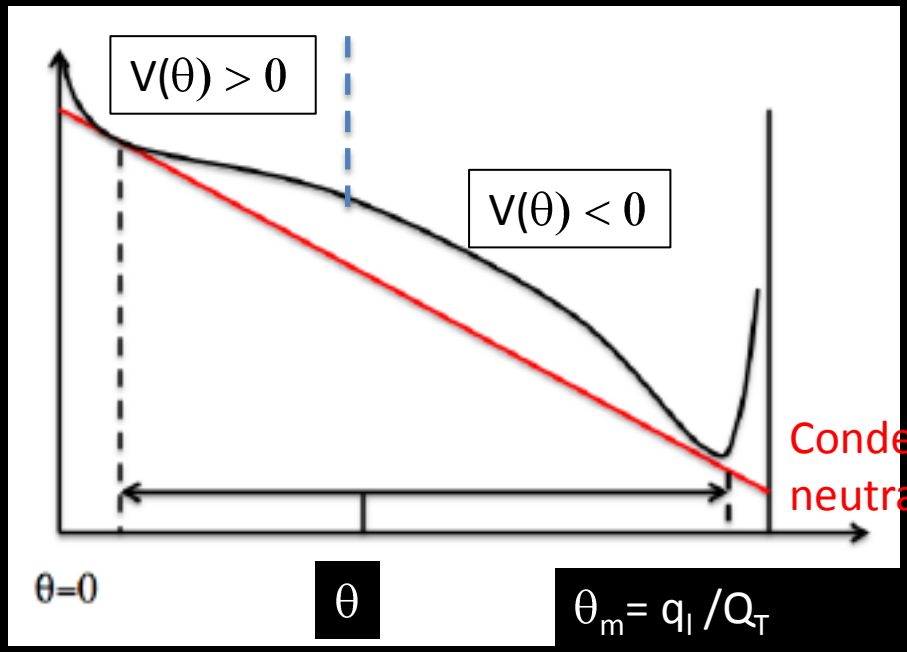
RNA/tail affinity

$$F_{RNA}(R,q) = F_F(R,q) + k_B T\, Nq \ln q - (m+e)Nq$$

Chemical potential tails

* Minimize with respect to R

Swollen, charged



Condensed, neutralized

- Common-tangent Construction: Phase Coexistence

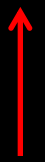
Large, reversible first-order phase transition

Gel swelling/shrinking

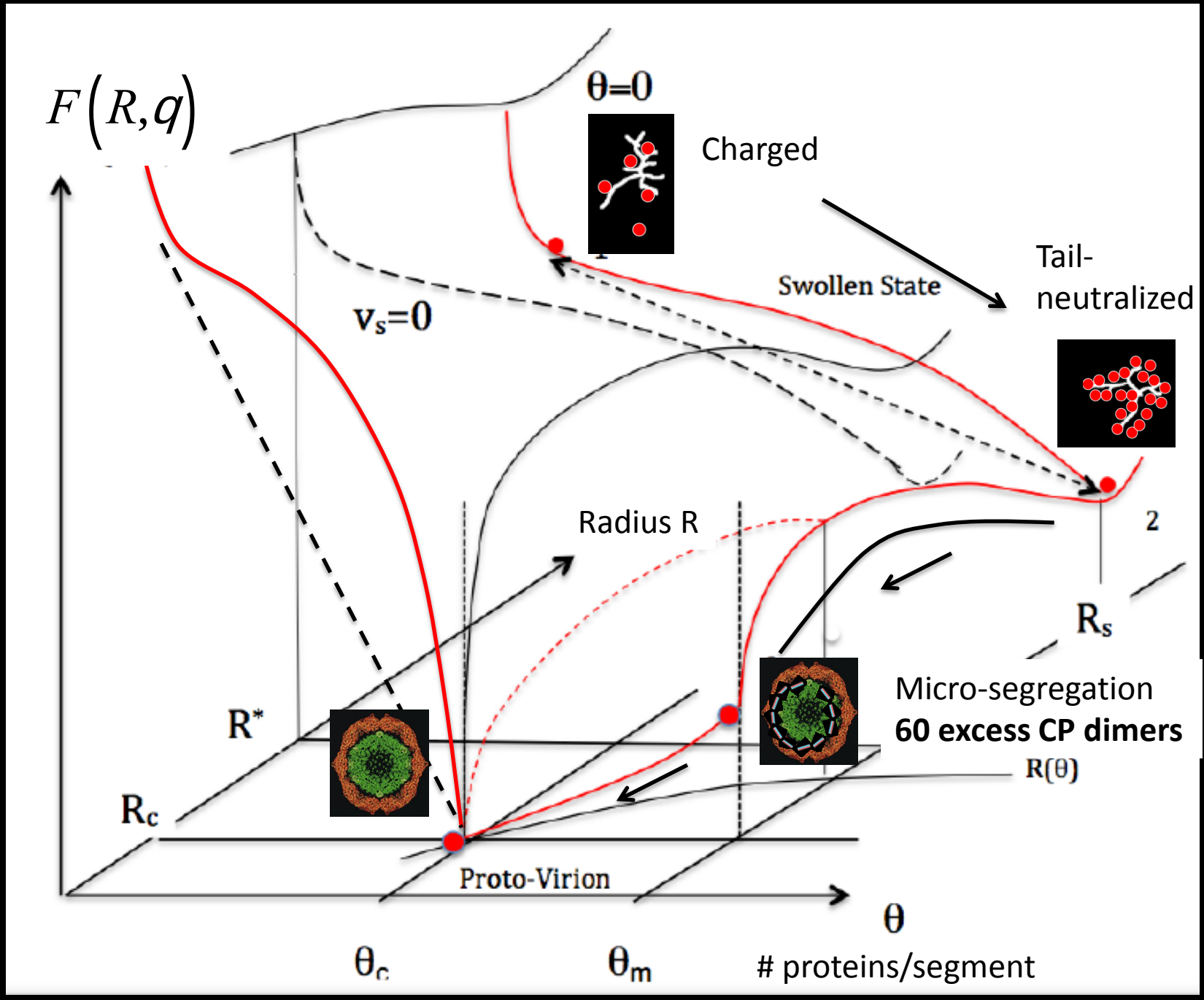
V) Free Energy Landscape

Combine:

$$F(R,q) = F_{RNA}(R,q) + F_S(R,Nq)$$

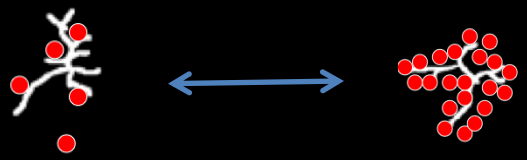


Surface-Adsorbed CPs = # Tails


$F(R,q)$
θ=0
Charged
Swollen State
Tail-
neutralized
v_s=0
Radius R
2
R_s
R^*
R_c
Micro-segregation
60 excess CP dimers
R(θ)
Proto-Virion
θ
θ_c
θ_m
proteins/segment

Is this processes thermodynamic reversible self-assembly?

Step 1

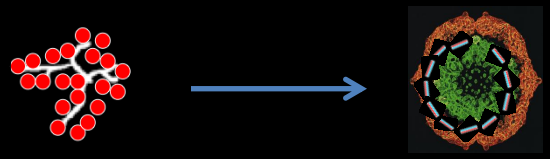


Reversible

Protein-RNA assembly

Same CP chemical potentials

Step 2



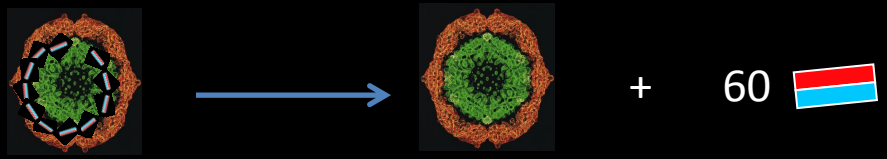
Irreversible

Micro-segregation

Lowered CP chemical potential
Enhanced RNA self-attraction

$$k_B T\left(\frac{q_B Q_C^2}{k} - J_S\left(R_c / D\right)\right) r_s$$

Step 3



Irreversible

Protein expulsion

Lowered CP chemical potential

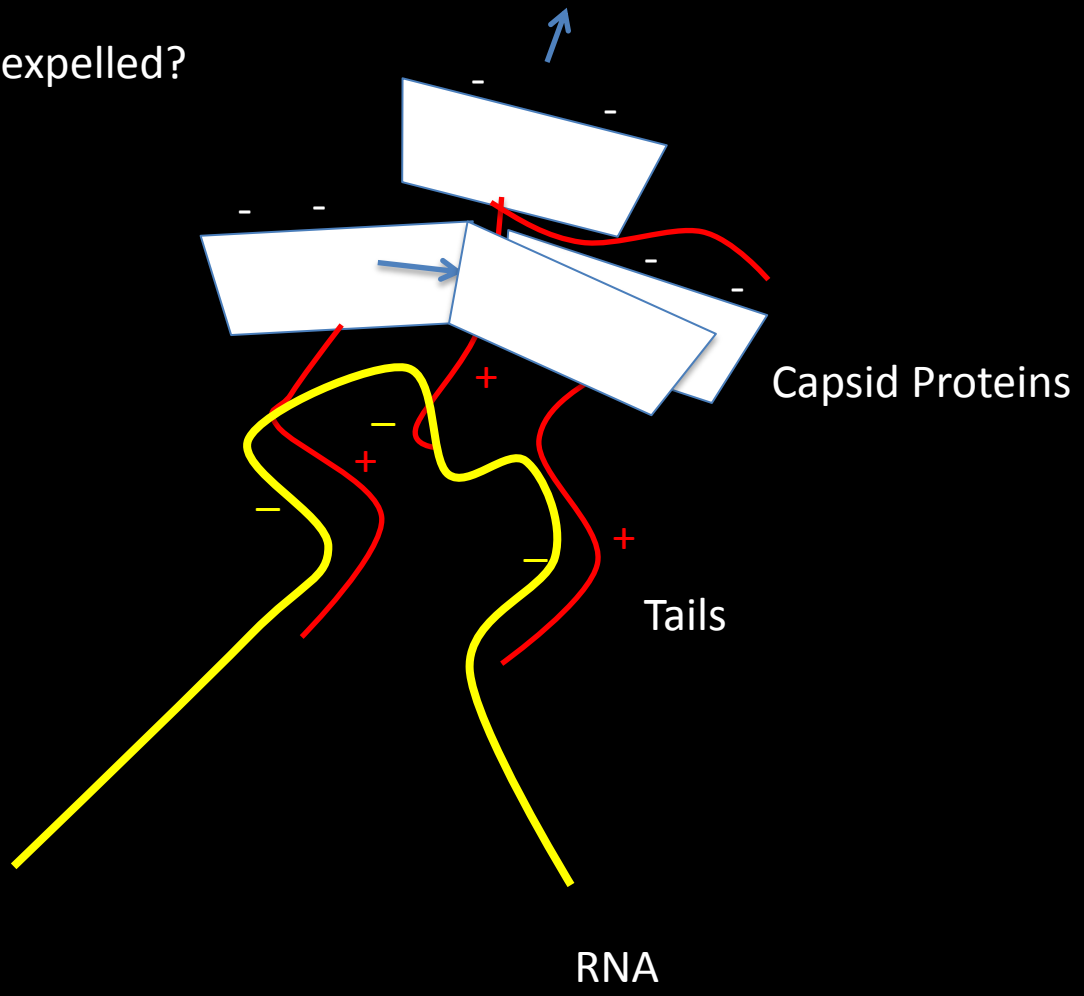
$$e^{-D V_D Q_C}$$

Donnan Potential + Protein Self-repulsion

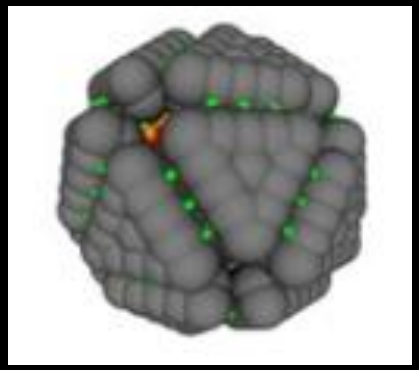
"Michaelis-Menten like"

How are excess proteins expelled?

Brownian Ratchet:

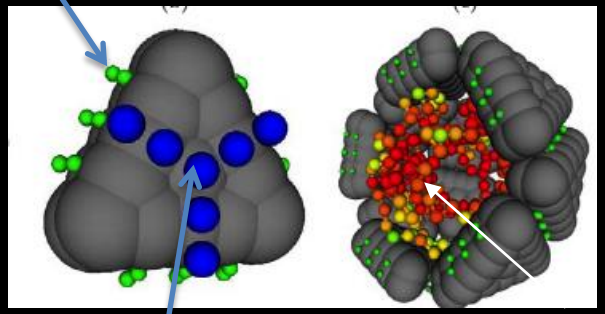


How good is mean-field theory?



Toy T=1 Virus

Protein-Protein binding sites



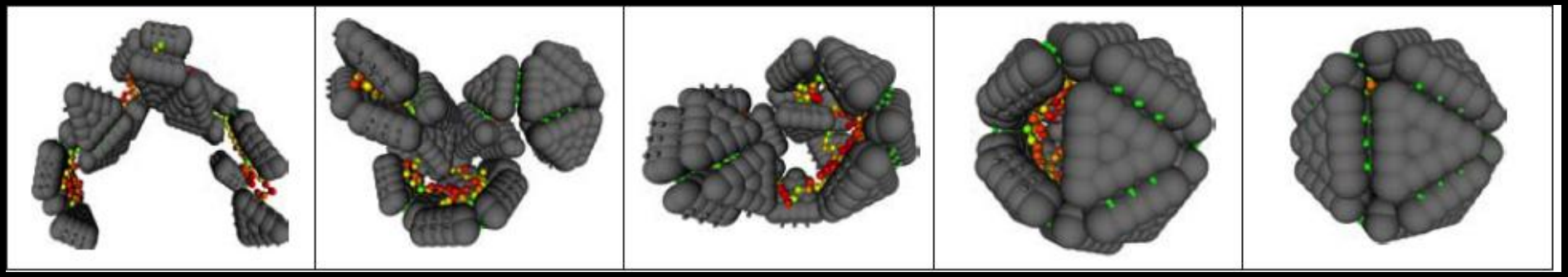
Flexible *linear* polymer genome

genome binding sites”

* Genome molecule: no branching.
* Assembled state: # binding sites = chain length

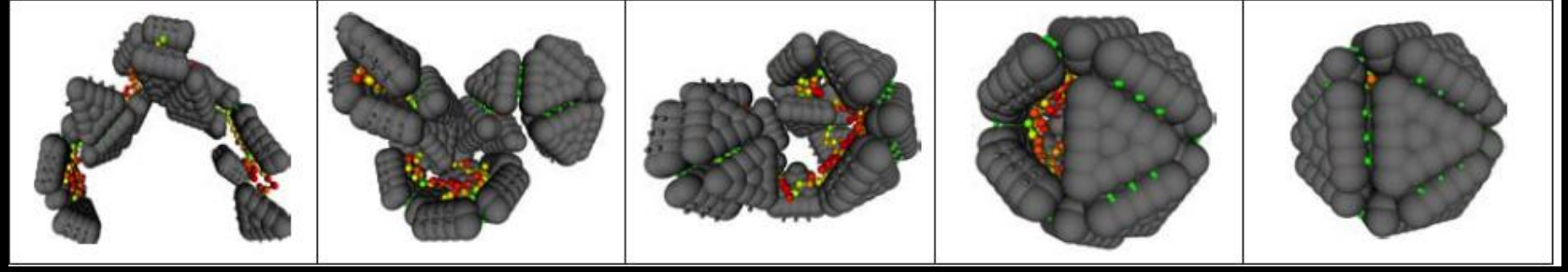
Elrad and Hagen

Protein-genome affinity ε > Protein-protein affinity J



time

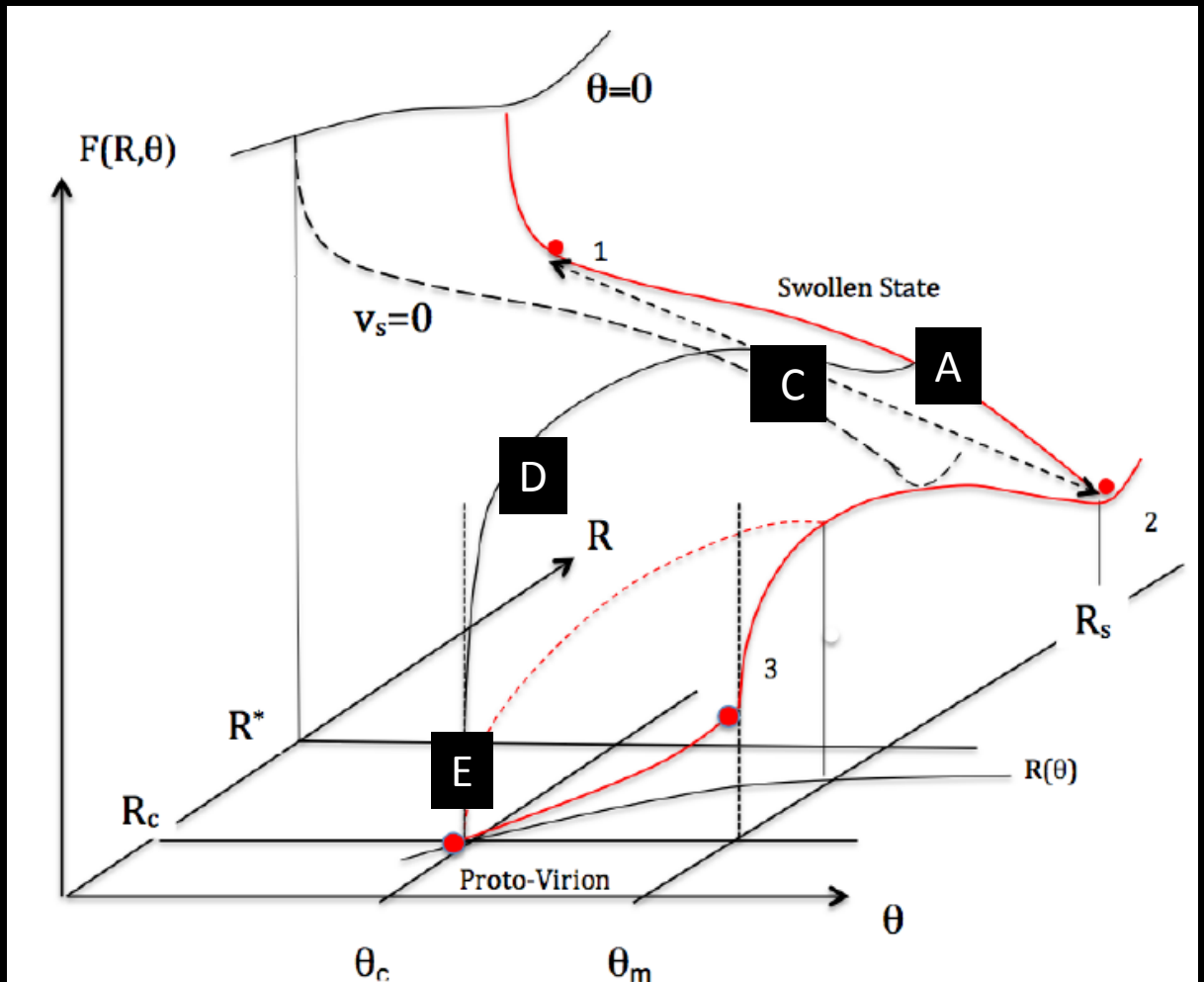
* RNA/Protein pre-assembly condensate



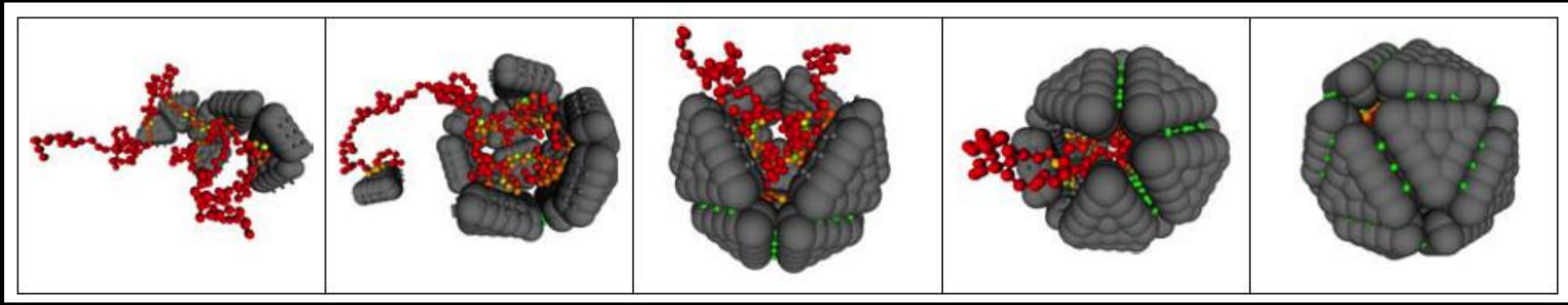
A B C D E

Problem: Optimal angles visible in A-C

- Local correlations.

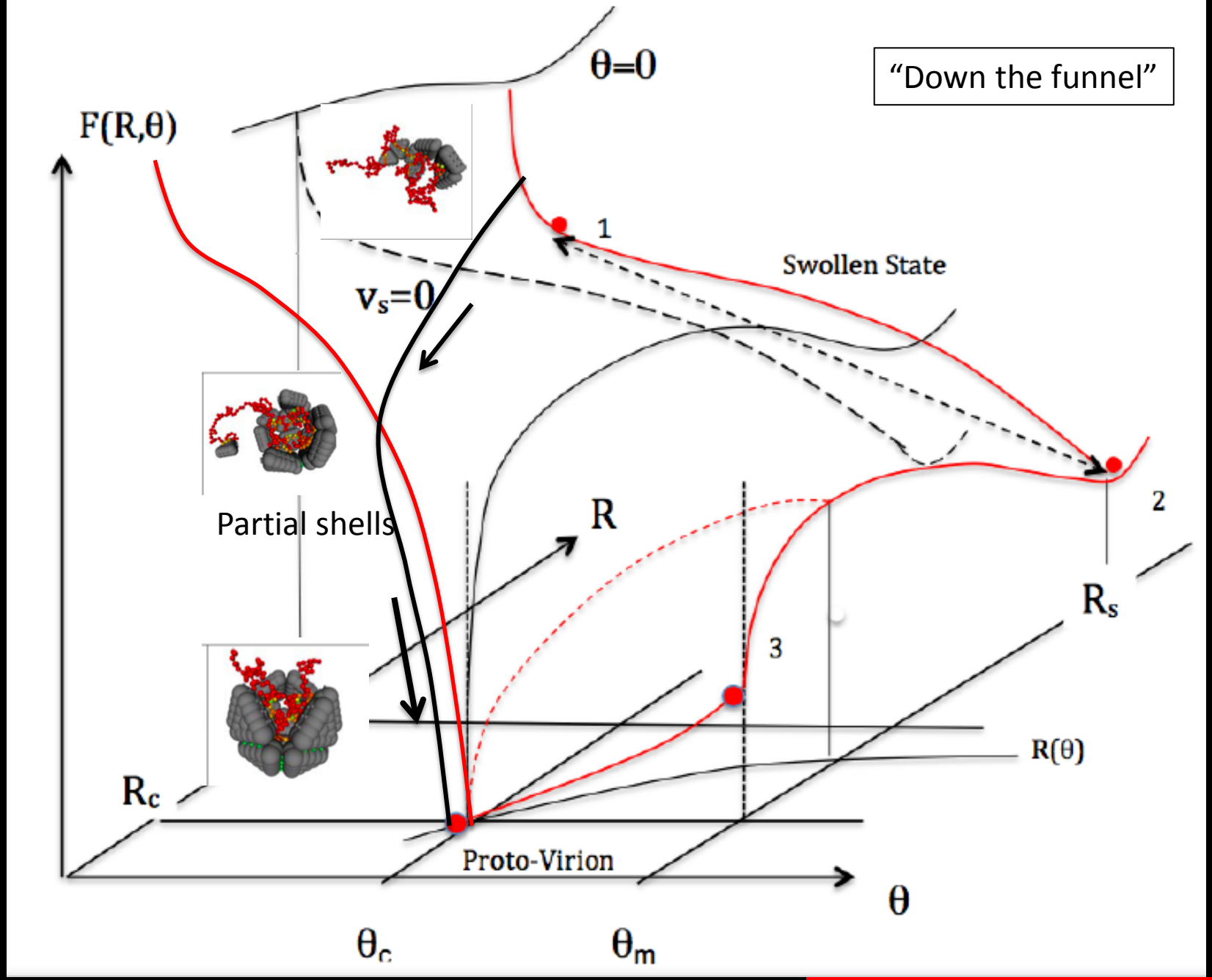


Genome-protein affinity ε weaker than protein-protein affinity J



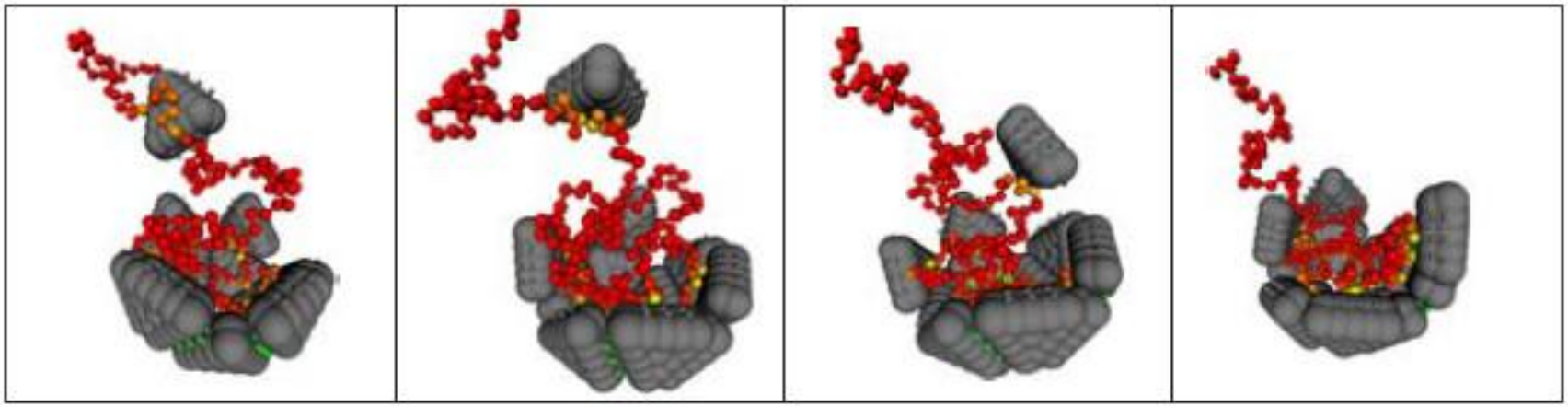
RNA “glues” capsomers together one-by-one

* *Heterogeneous nucleation* of a shell on a flexible RNA scaffold



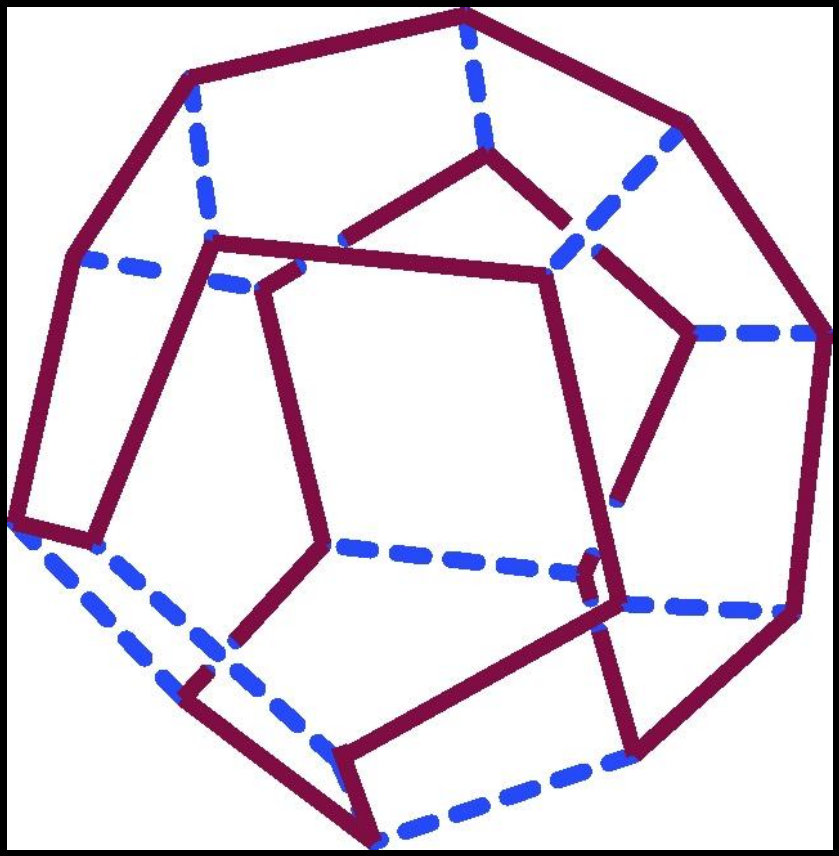
Many possible assembly pathways

“Antenna-Assembly”



(Hu and Shklovskii)

“Hamiltonian Cycle”



- Graph-theoretic problem (R.Twarock)

Conclusions

1) Assembly of small ss RNA viruses can be viewed as the combination of reversible RNA condensation + quasi-reversible shell formation.

2) Combination of two simple thermodynamic assembly processes produces a more complex free energy landscape with different possible multi-step Irreversible pathways.

3) Viral assembly appears intermediate between Mark I and Mark II assembly.