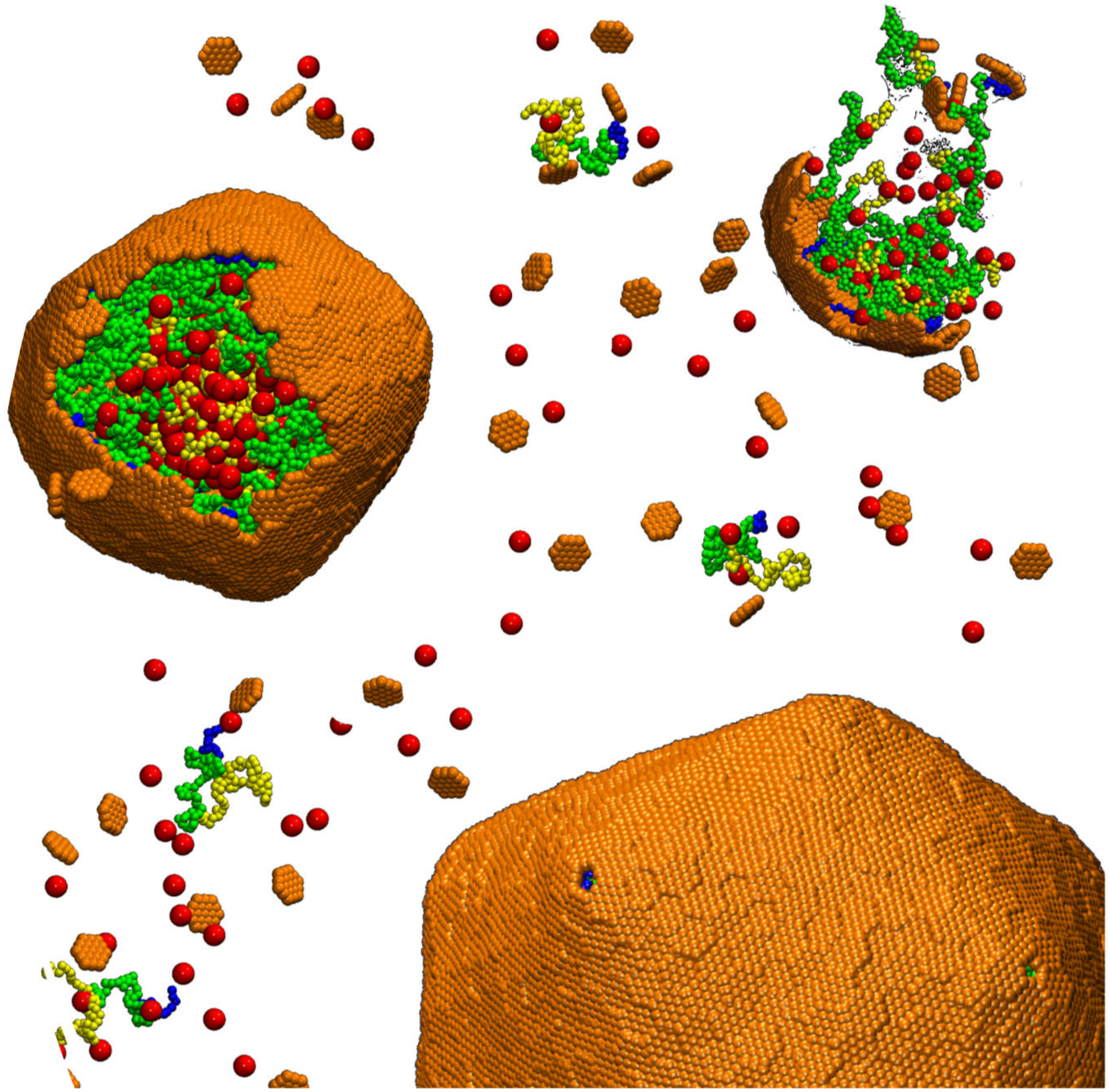
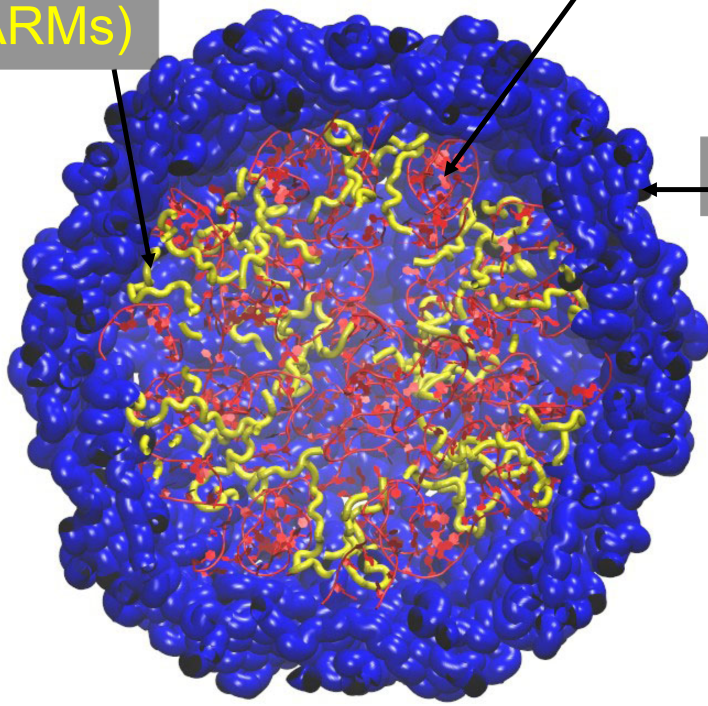
Computational modeling of microcompartment assembly and cargo condensation

Michael F. Hagan
Farri Mohajerani
Lev Tsildokovski
Department of Physics and Quantitative Biology Program, Brandeis University

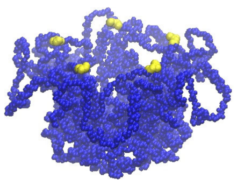


Electrostatics help drive protein-RNA association

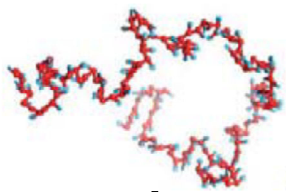


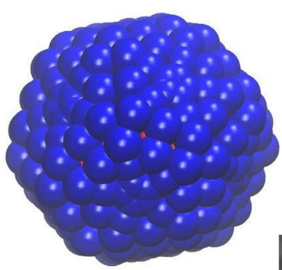
STMV [coordinates from
Y. Zeng … S. Harvey,
JMB, **180** (2012) 110–116]

Experiments on Capsid Assembly around ssRNA

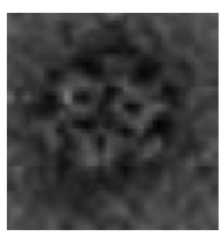
12  SV40 subunit

\+ ssRNA

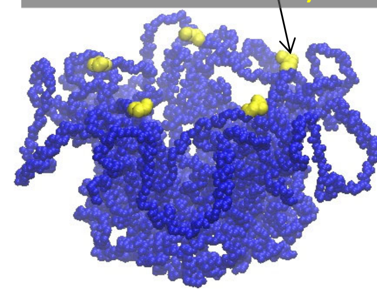




Dodecahedron virus-like particle



ARMs (not resolved in structure)



SV40 subunit

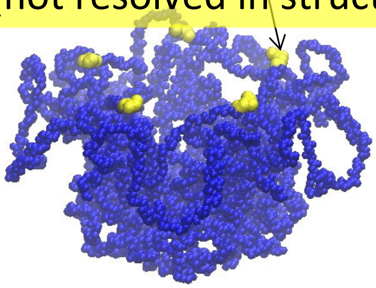
Kler, …,Oppenheim, Zlotnick, Raviv JACS (2012), ACS Chem. Biol. (2013)

Model for ssRNA Viruses

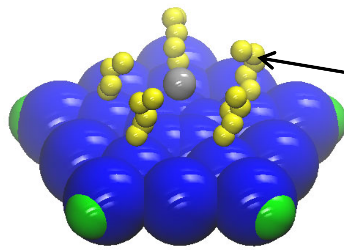
Jason Perlmutter

ARMs (not resolved in structure)

SV40 subunit

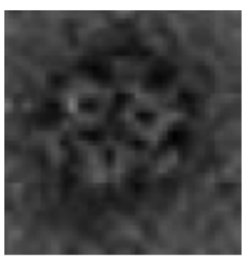


model subunit



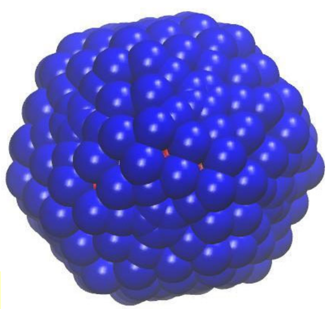
yellow = +1 charge

SV40 capsid



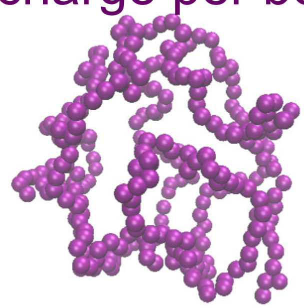
Kler, et al JACS (2012), ACS Chem Biol (2013)

model capsid



linear polyelectrolyte

-1 charge per bead

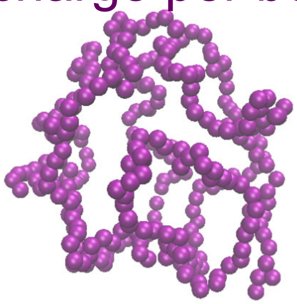


Perlmutter, Qiao, MFH, ELife (2013)
Perlmutter, Perkett, MFH, JMB (2014)
Perlmutter & MFH JMB (2015)

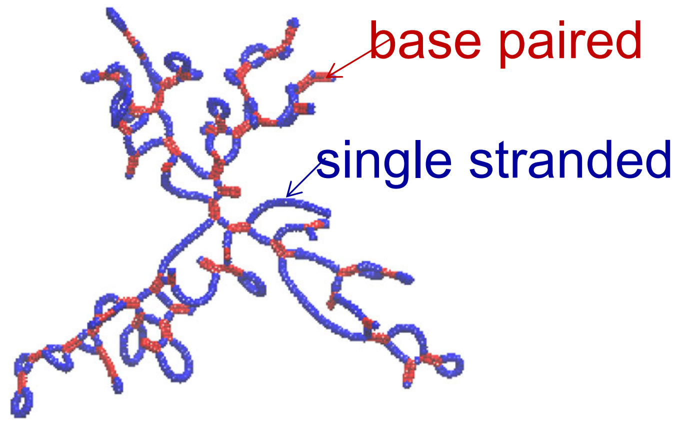
Representing RNA Structure

linear polyelectrolyte

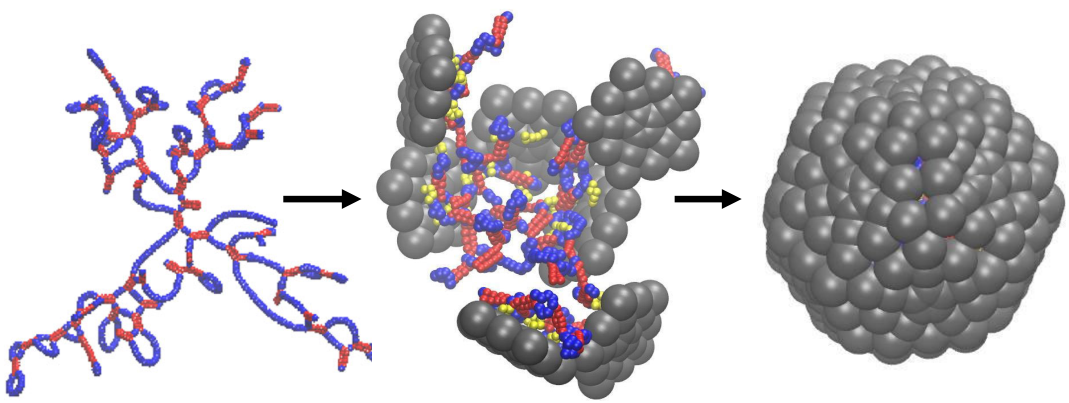
-1 charge per bead



ssRNA



secondary structure is static (for now)



Perlmutter, Qiao, MFH, ELife (2013)
Perlmutter, Perkett, MFH, JMB (2014)
Perlmutter & MFH JMB (2015)

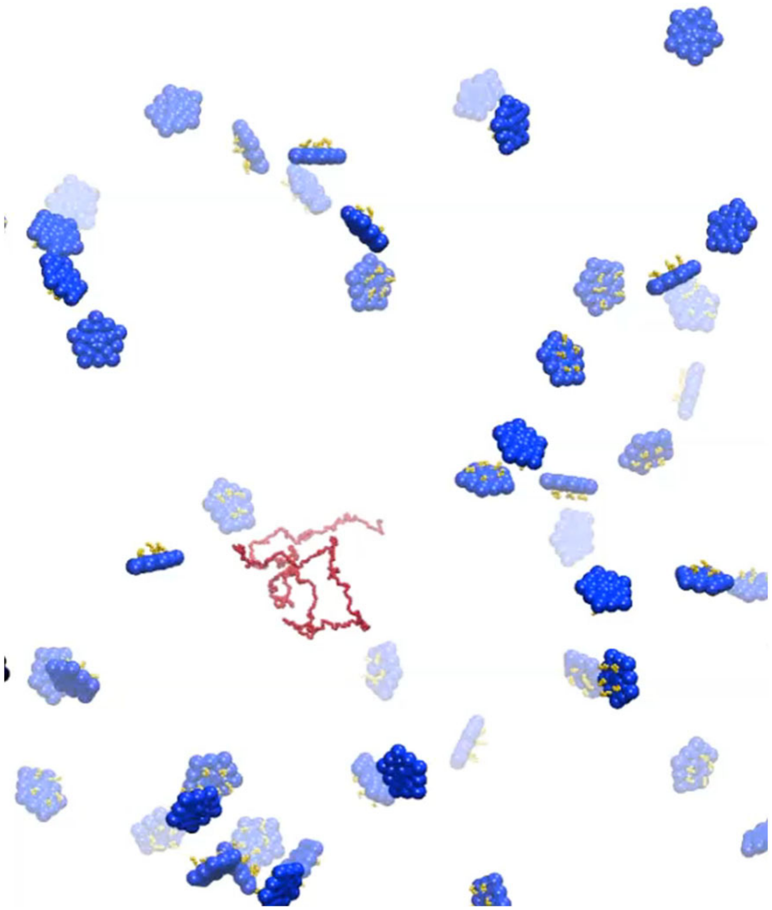
Example assembly trajectory

polymer
charge=-600*e*

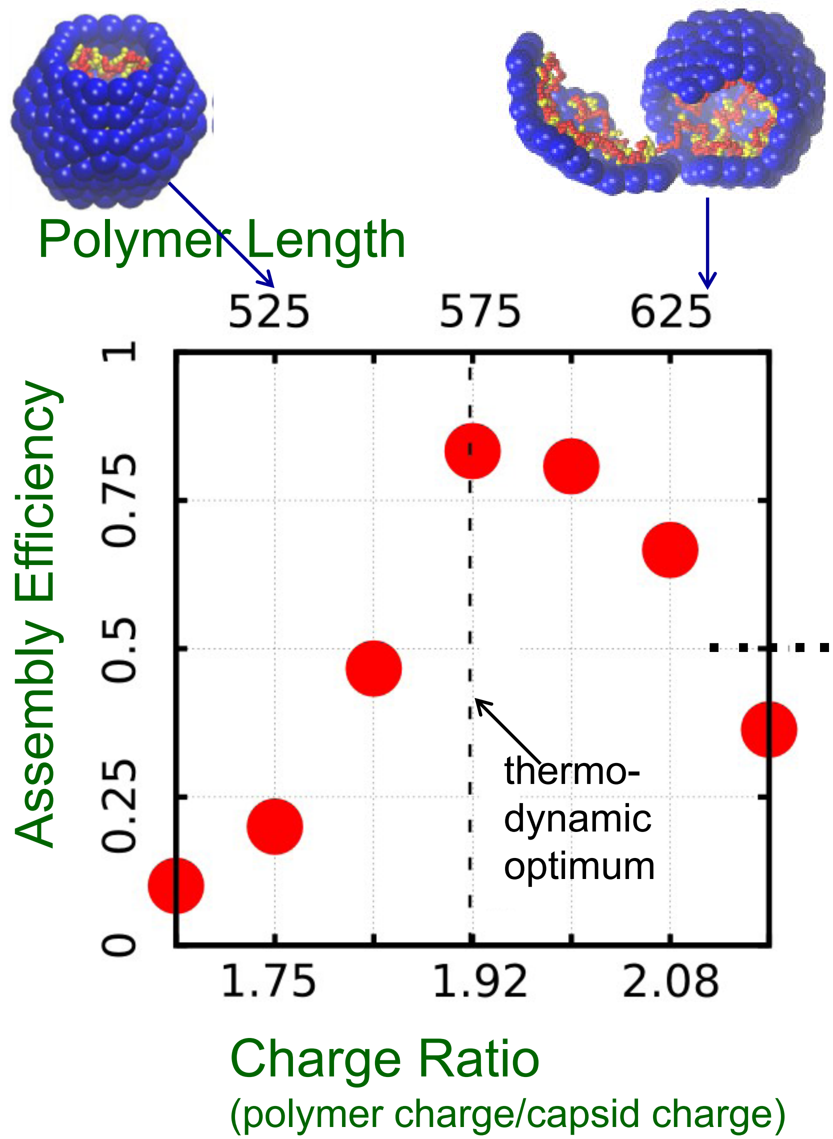
capsid
charge=300*e*

[salt] = 100 mM

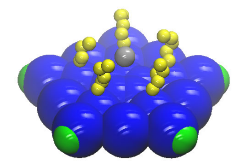
simulated on GPUs using
HOOMD



Perlmutter, Qiao, MFH, *eLife*, 2:e00632 (2013)

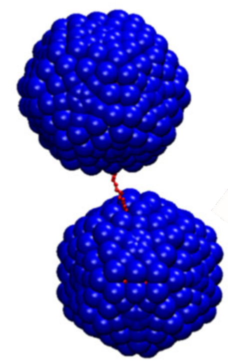


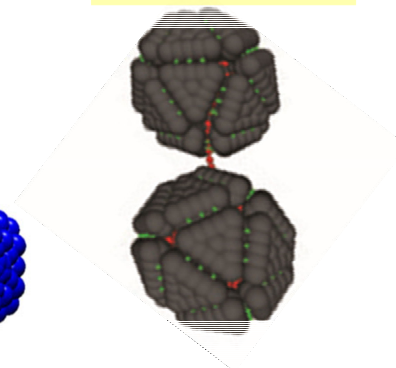
capsid charge =
300e

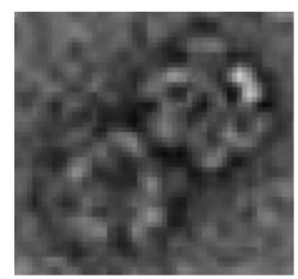


1200

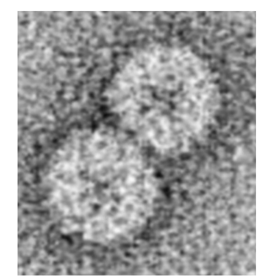
Elrad & MFH
(2010)







Kler et. al. ACS
Chem. Biol (2013)



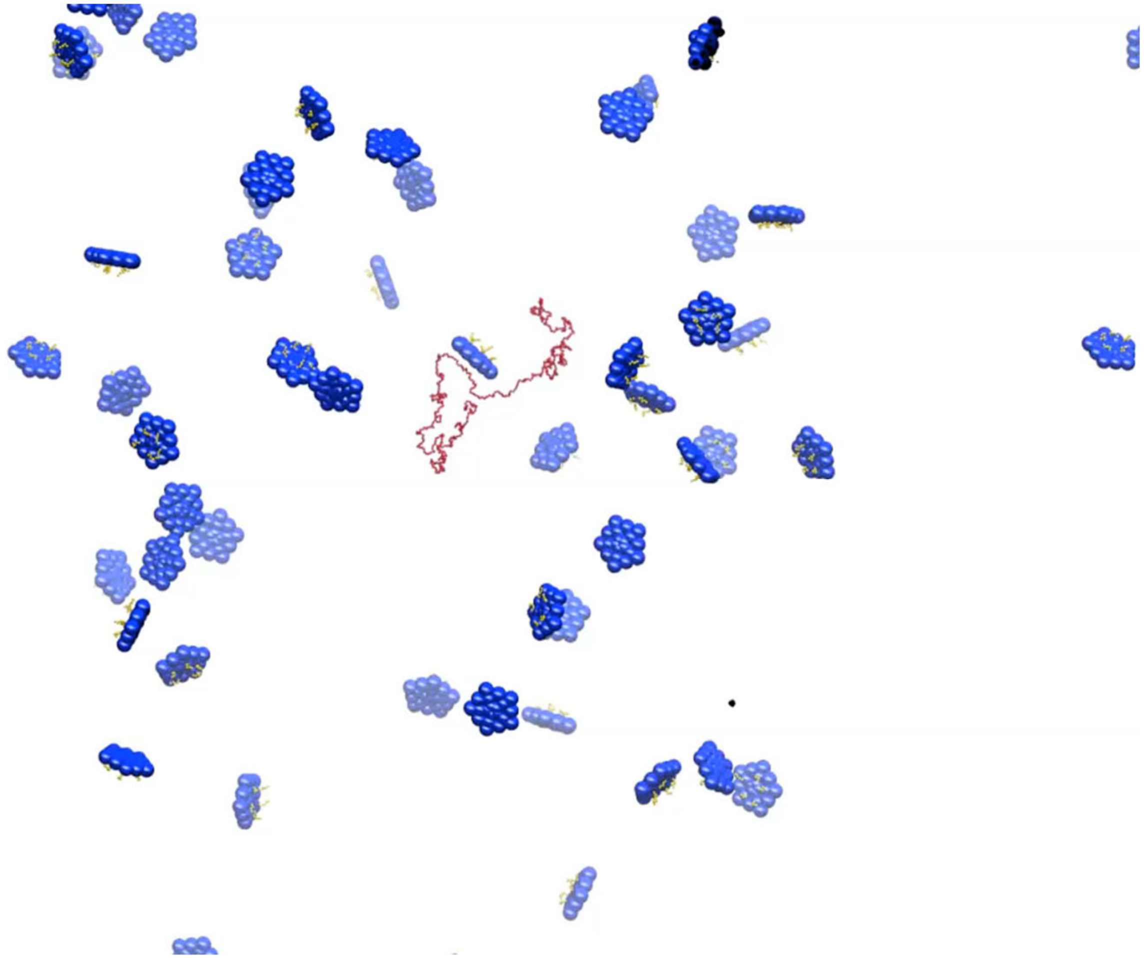
Cadena-Nava et. al.
J. Virol (2012)

Assembly Trajectory at Higher Salt

polymer
length=600

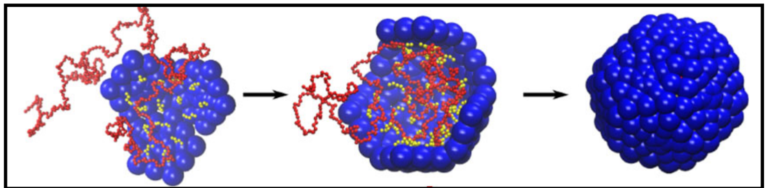
capsid
charge=300

[salt] = 300 mM

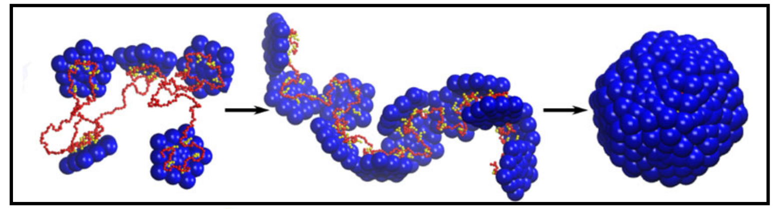


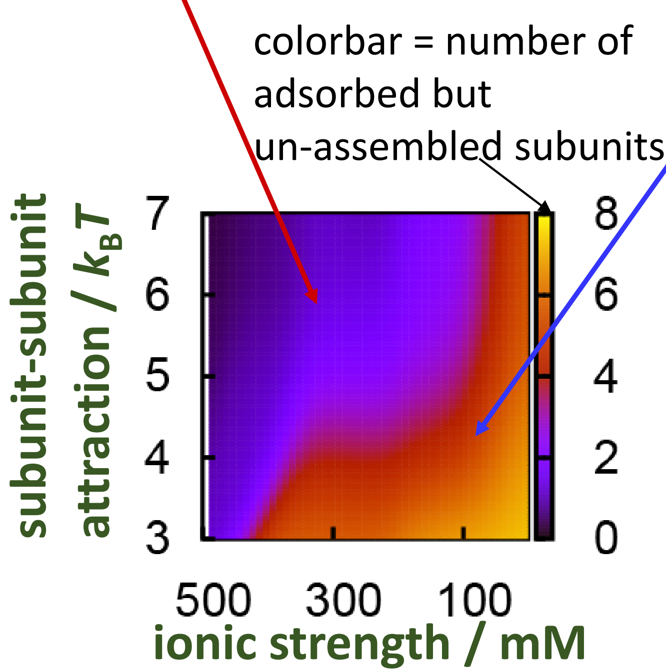
Controlling Assembly Pathways

nucleation-and-growth (NG)



en masse





increasing protein-protein interactions

increasing protein-polyelectrolyte interactions

Experiments:
Garmann et al. JMB (2014);
Garmann et al. Act. Chem. Res. (2015)

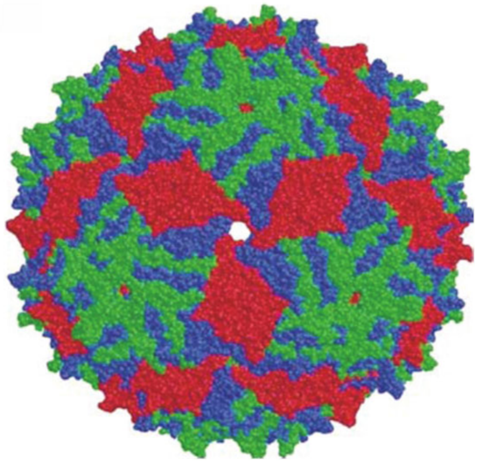
Garmann et. al, bioRxiv (2022)
https://doi.org/10.1101/2022.05.09.488235

Perlmutter, Perkett, MFH,
JMB 2014; Perlmutter & MFH JMB 2015

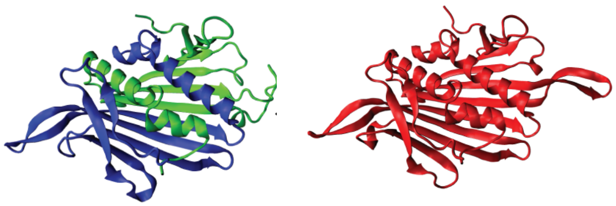
Model for MS2 Assembly

Walter Singaram

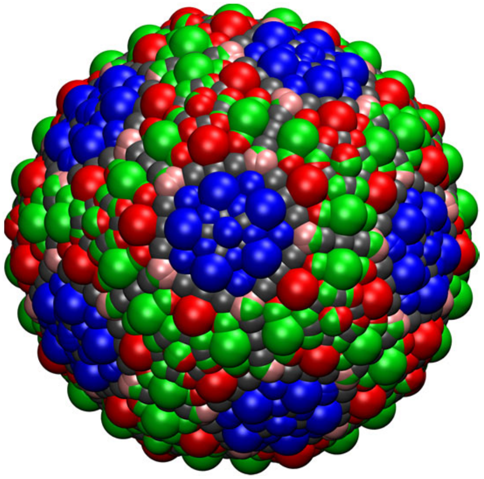
MS2 capsid, 90 protein dimers (T=3)



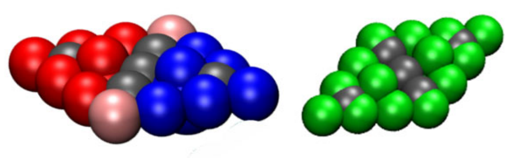
protein dimers



Model capsid

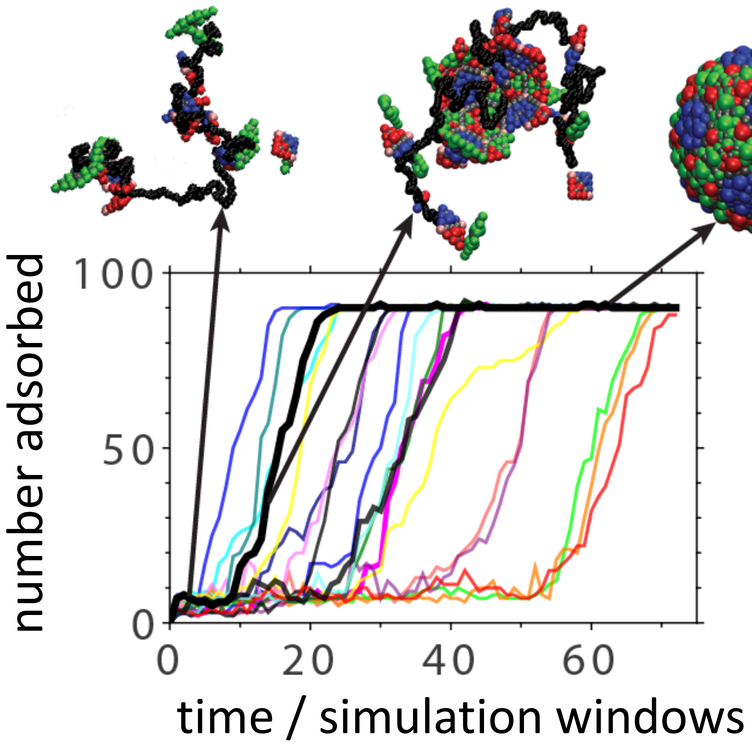


model protein dimers

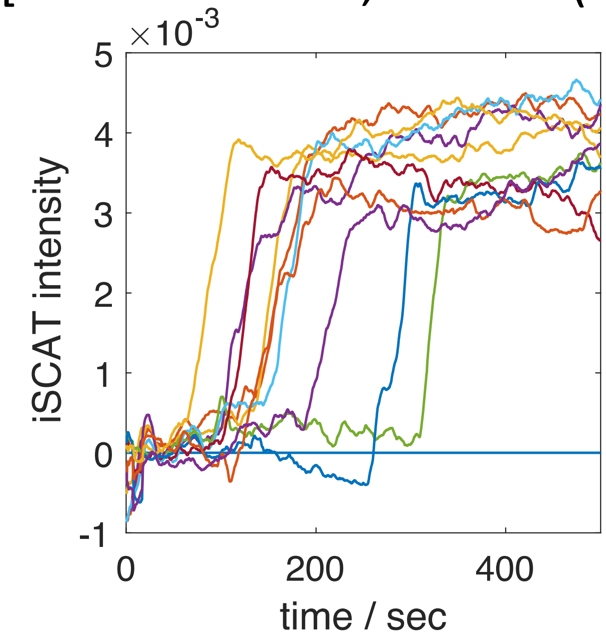


Simulations of MS2 Assembly

Individual simulation trajectories for parameters leading to nucleation-and-growth pathways

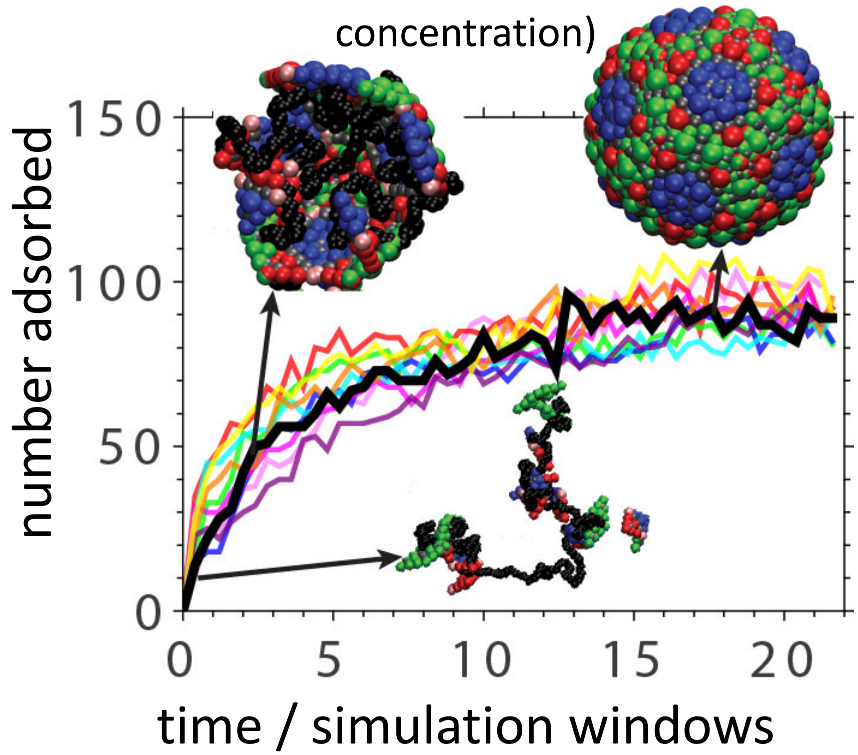


experiments monitoring assembly of individual capsids [Garmann et. al, bioRxiv (2022)]



Simulations of MS2 Assembly

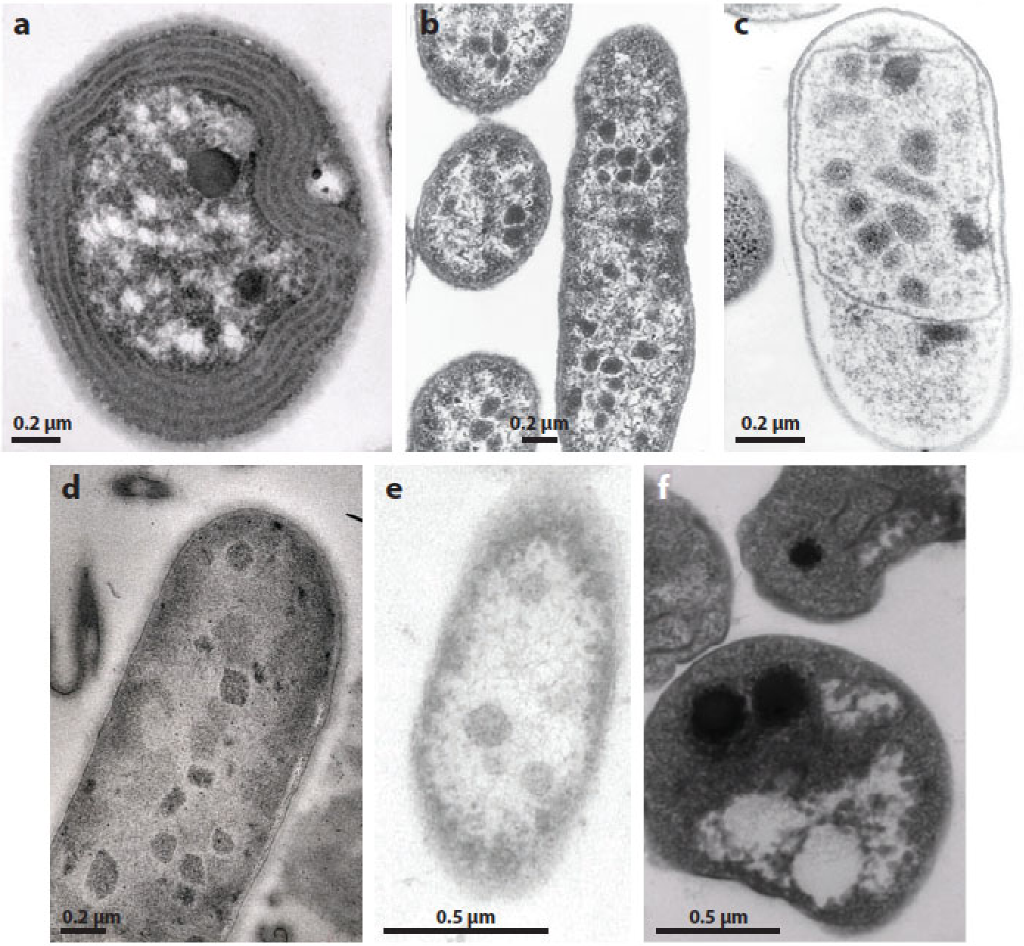
simulation trajectories for parameters leading to growth limited (higher binding energy, lower concentration)



Bacterial Microcompartments (BMCs)

- BMCs are protein-shelled compartments in bacteria
- BMCs sequester and/or concentrate reactants

TEM images of various BMCs

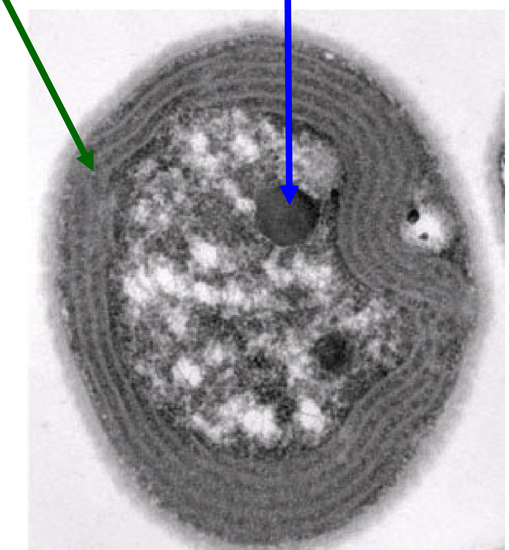


Kerfeld et al. Annu. Rev. Microbiol. 2010

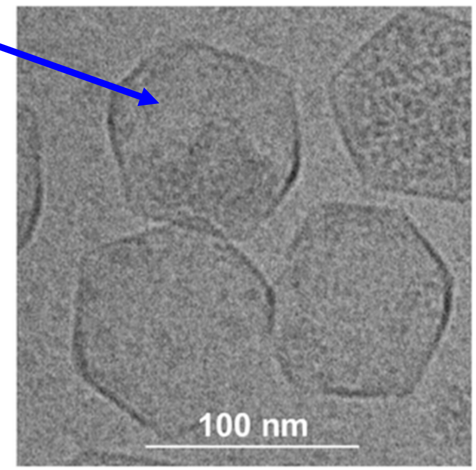
Bacterial Microcompartments

cyanobacteria ("blue-green algae")

carboxysome



1 μm

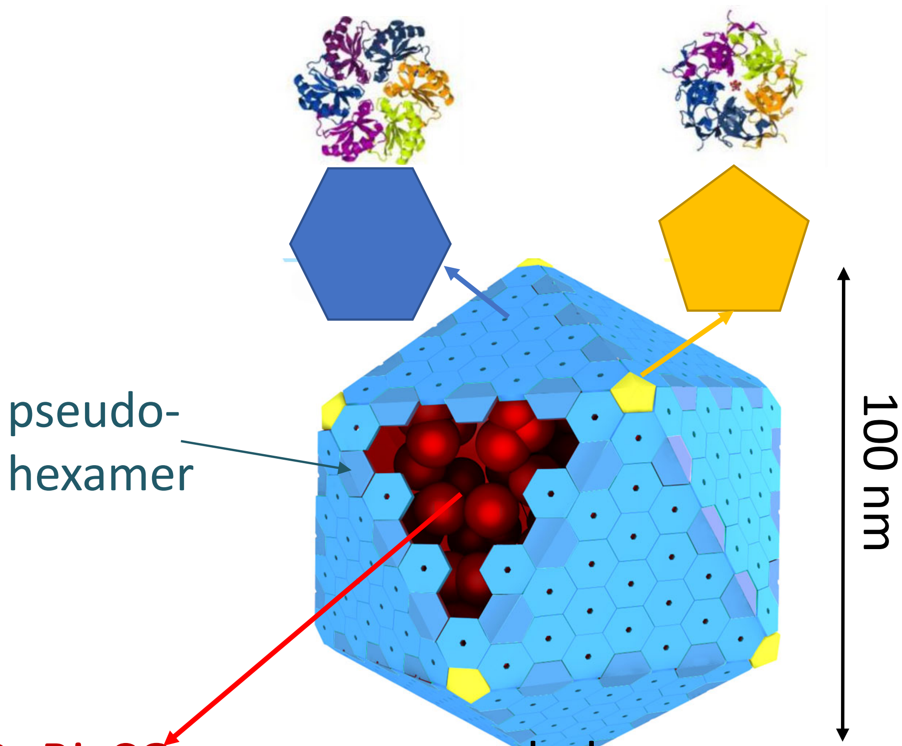


Wikipedia

Kerfeld et al. Annu. Rev. Microbiol. 2010

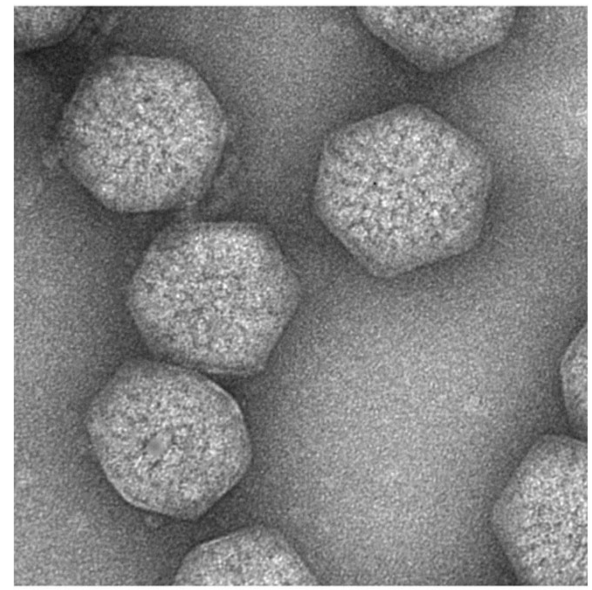
Carboxysome structure

Hexameric and Pentameric shell proteins



RuBisCO = enzyme needed to convert photosynthesis products to sugars

Kinney … Kerfeld, Photosynth Res (2011), Tanaka, Kerfeld, … Yeates, Science (2008)

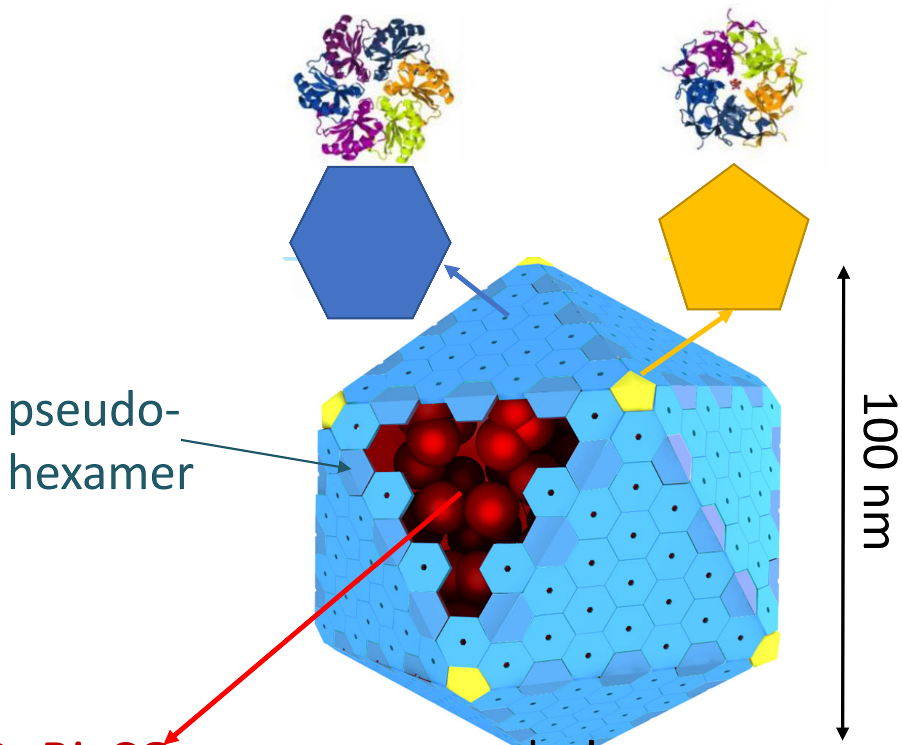


Carboxysome in Cyanobium cyanobacteria

Long (2018) Nature Communications

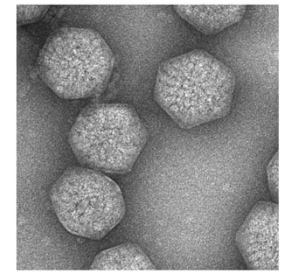
Carboxysome structure

Hexameric and Pentameric shell proteins



RuBisCO = enzyme needed to convert photosynthesis products to sugars

Kinney … Kerfeld, Photosynth Res (2011), Tanaka, Kerfeld, … Yeates, Science (2008)



Long et al. (2018) Nature Communications

re-engineering microcompartments as customizable nanoreactors

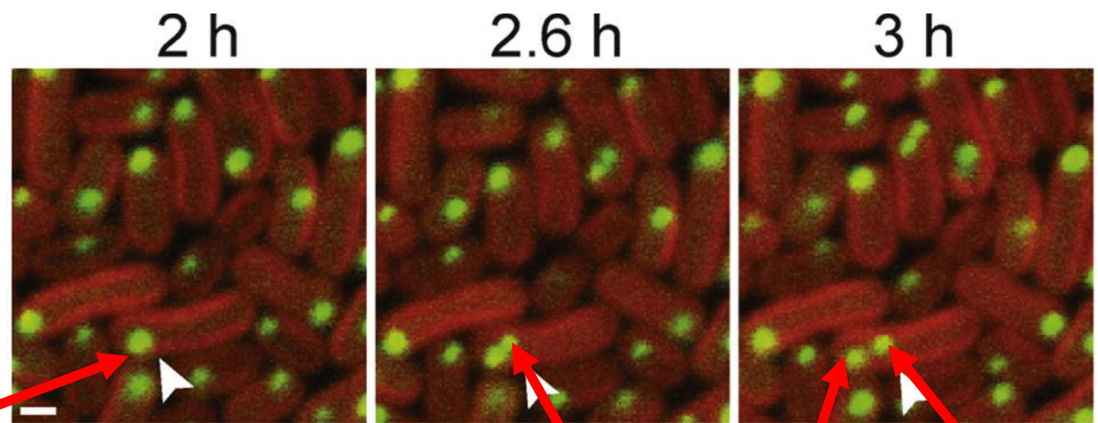
[e.g. Lee and Tullman-Ercek Curr. Opin. Sys. Biol. (2017)] https://doi.org/10.1016/j.coisb.2017.05.017, C. A. Kerfeld et al., Nat. Rev. Microbiol. 16, 277 (2018)]

what factors control:

- size of shell
- amount of cargo
- composition of cargo

Carboxysome Assembly in Bacteria

RuBisCO
labeled green



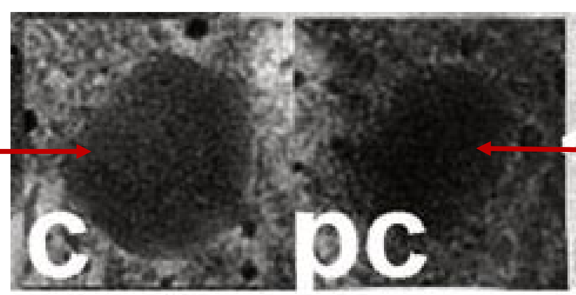
Procarboxysome
(dense RuBisCO
complex)

Carboxysome budding
from procarboxysome

Carboxysome

Procarboxysome

Carboxysome

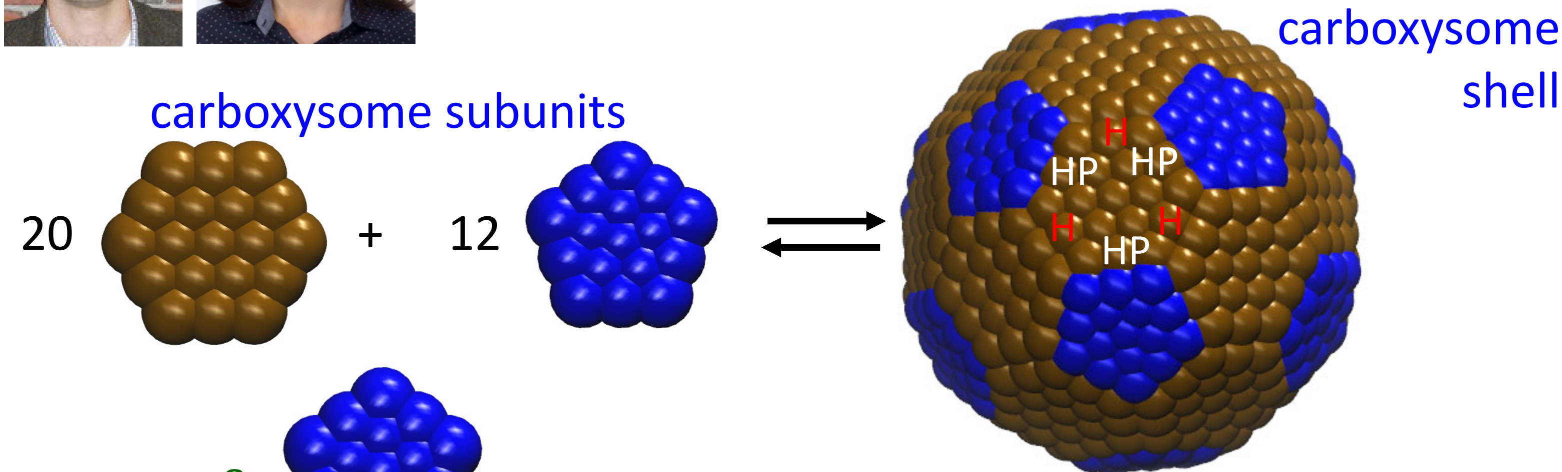


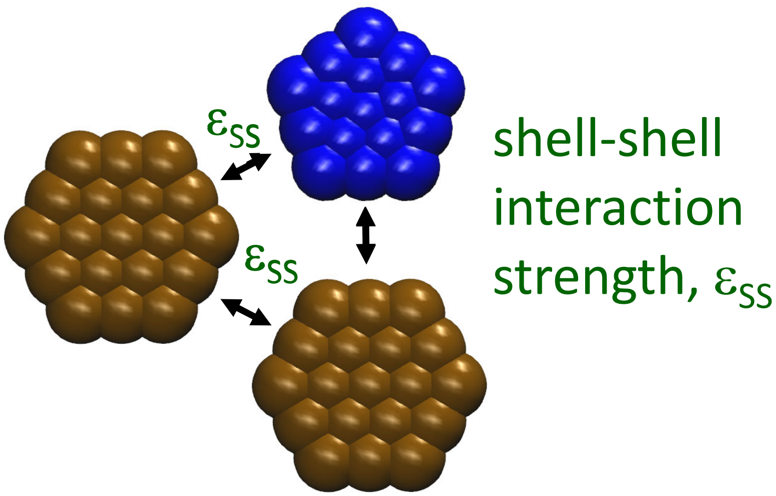
Procarboxysome

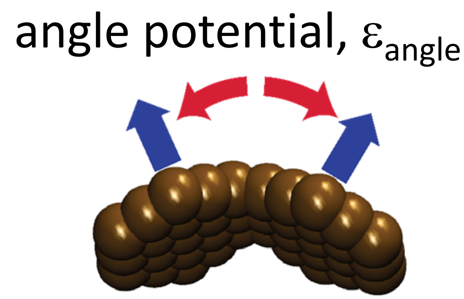
Cameron, Wilson, Bernstein, Kerfeld, Cell (2013)

Model: Carboxysome Shell

Jason Perlmutter, Farri Mohajerani





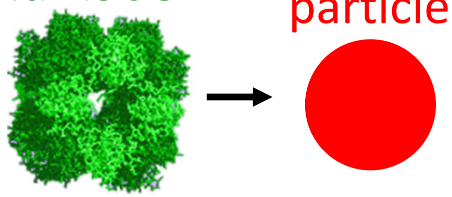


Perlmutter, Mohajerani, MFH, ELife (2016), Mohajerani&MFH, Plos. Comp. Biol. (2018), Mohajerani et. al ACS Nano (2021)
see also: Mahalik et al. ACS Nano (2016), Rotskoff&Geissler, PNAS (2018); S Li, DA Matoz-Fernandez, M Olvera de la Cruz ACS Nano (2021)

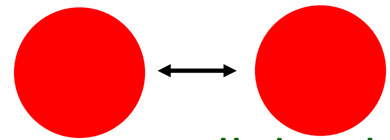
Model: Cargo

RuBisCO

Model cargo particle

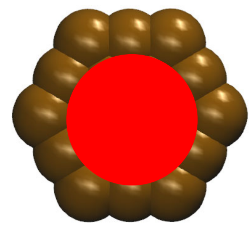


Lennard-Jones interaction



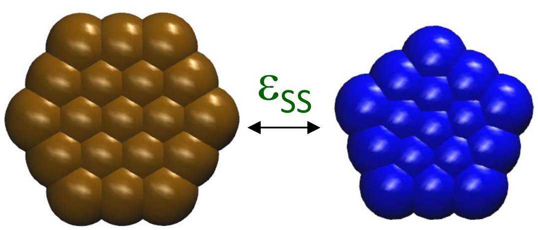
Lennard-Jones well depth, ε_{CC}

Cargo is attracted to inner surface of shell subunit

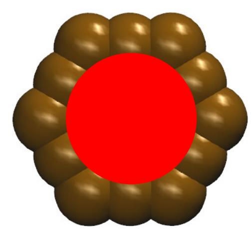


Shell-Cargo interaction strength, ε_{SC}

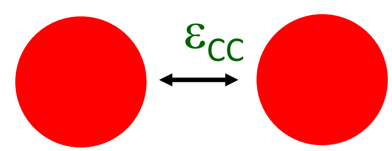
Interaction parameters



shell-shell, ε_{SS}



shell-cargo, ε_{SC}



cargo-cargo, ε_{CC}

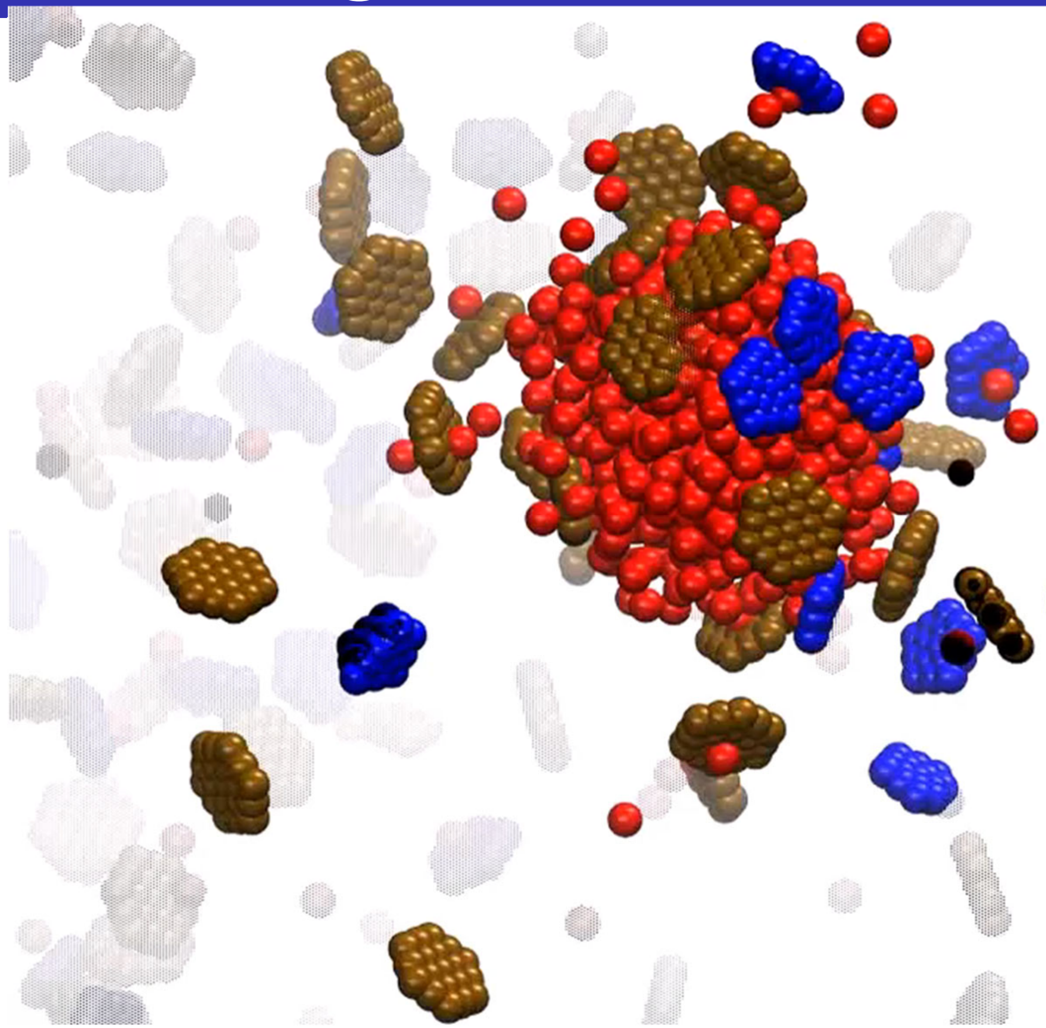
Perlmutter, Mohajerani, MFH, ELife (2016)

Assembly around a globule

shell-shell affinity
ε_{SS}=3.5 $k_B T$

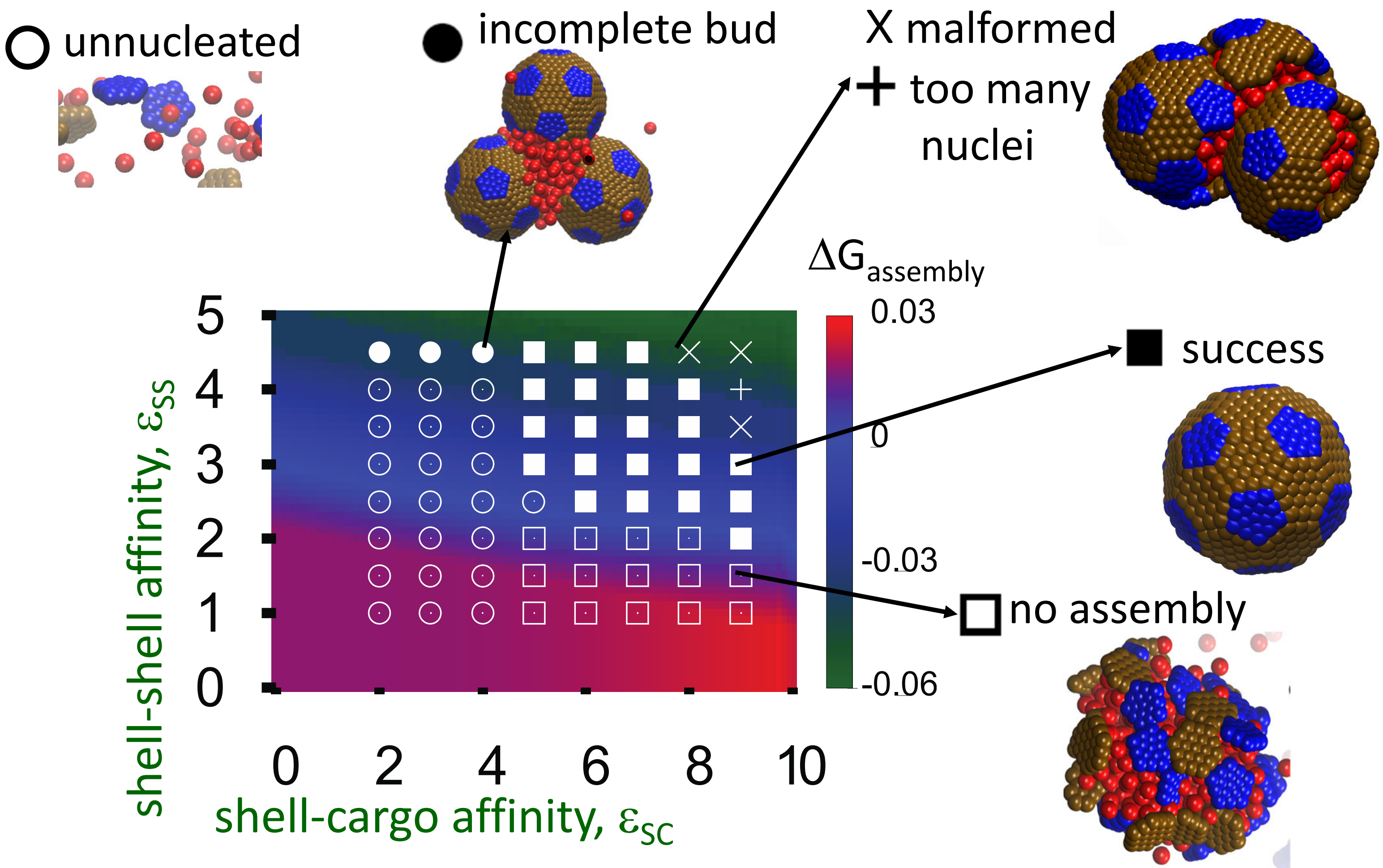
shell-cargo affinity
ε_{SC}=6 $k_B T$

cargo-cargo
affinity
ε_{CC}=1.6 $k_B T$



Perlmutter, Mohajerani, MFH, ELife (2016)

Phase diagram and assembly behavior

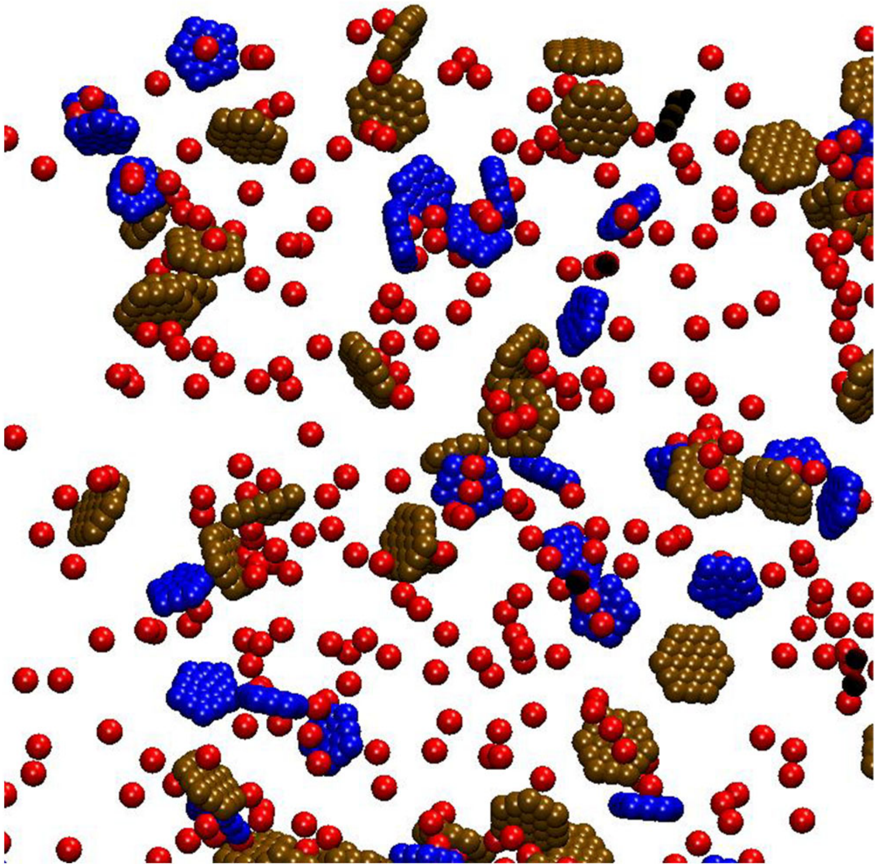


Perlmutter, Mohajerani, MFH, ELife (2016)
Mohajerani&MFH, Plos. Comp. Biol. (2018)

Cargo-Cargo affinity, ε_{CC}=1.6

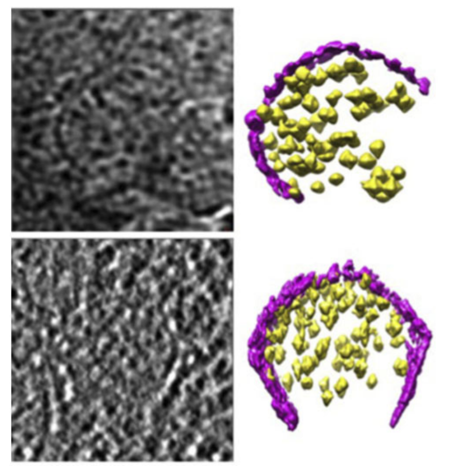
All energies in k_BT

Weaker cargo-cargo interactions lead to **simultaneous** assembly and cargo condensation



cargo-cargo affinity, ε_{CC}=1.3 $k_B T$

resembles
α-carboxysome
assembly

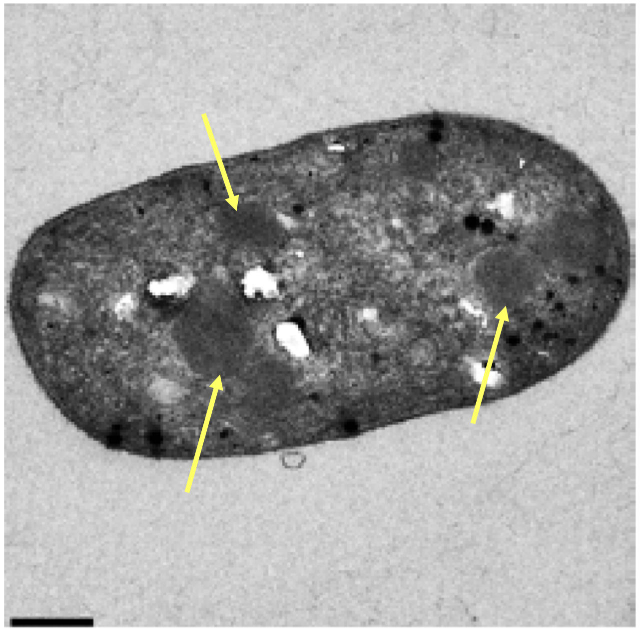


Dai et al., J. Mol. Biol. (2018)
Iancu et al, J. Mol. Biol. (2010)

Perlmutter, Mohajerani, MFH, ELife (2016)
Mohajerani&MFH, Plos. Comp. Biol. (2018)

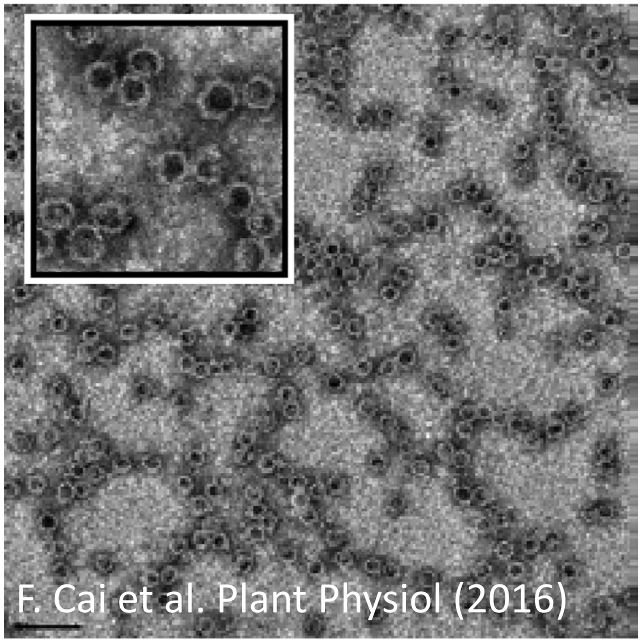
Shell Size

full shells assembled in cyanobacteria are **large** (~200 nm) and **polydisperse**



500 nm

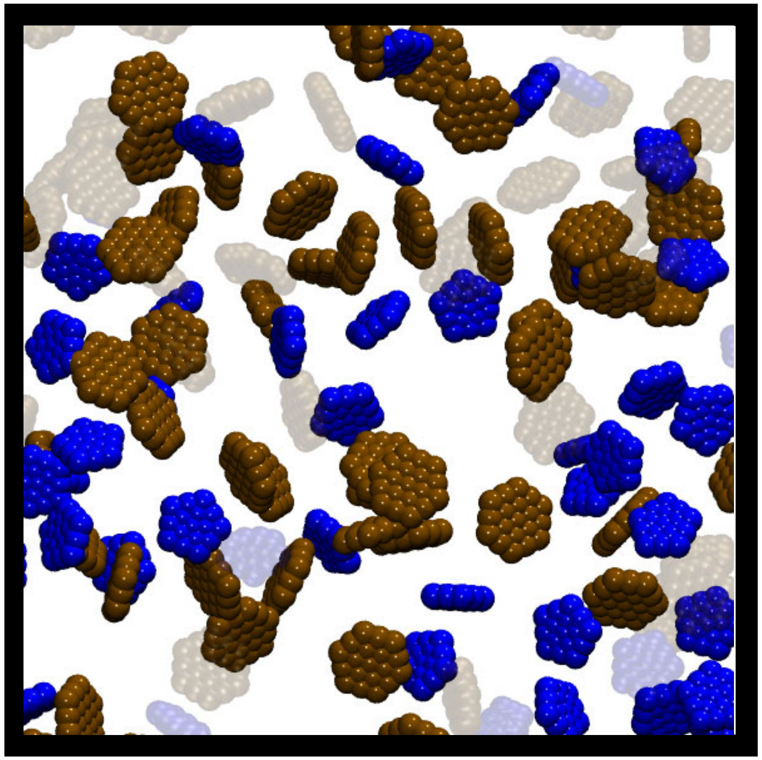
Empty shells assembled in E.Coli are **small** (20 nm) and **monodisperse**

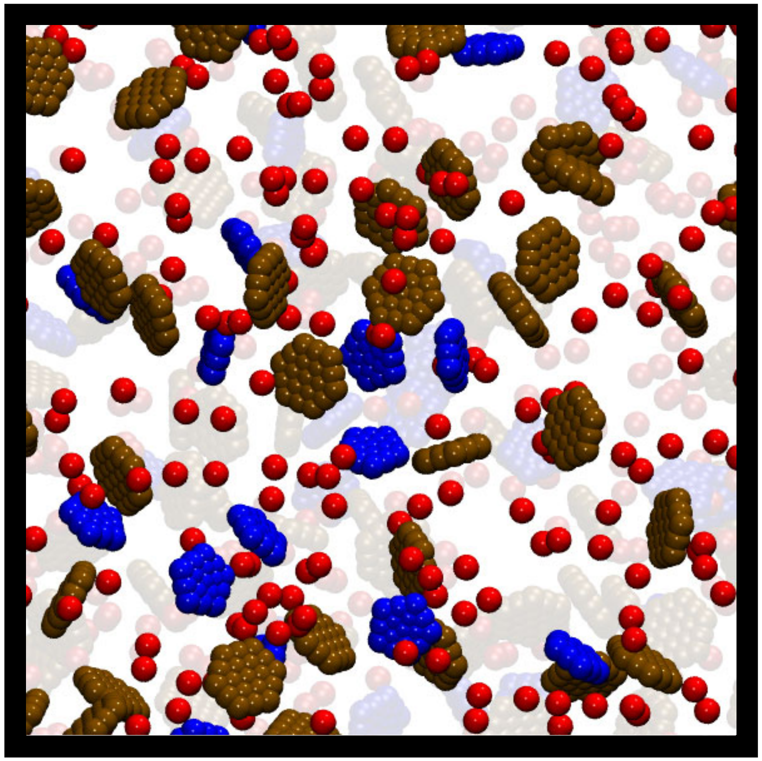


20 nm

Empty shells smaller than filled shells for other BMCs: Lehman et al., J. Bacteriol., 2017

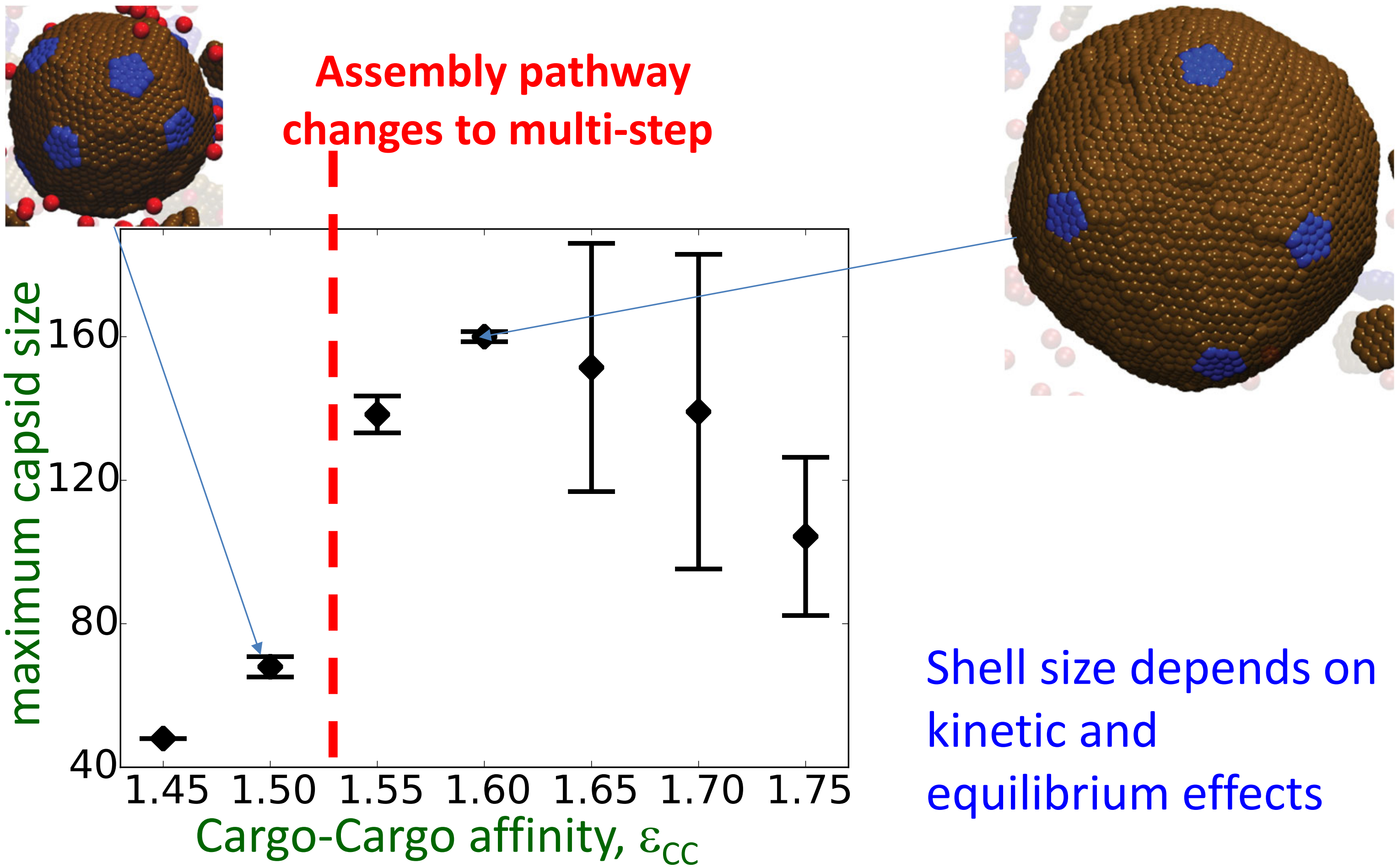
Cargo-controlled shell size





Only hexamers interact with cargo in microcompartments

Shell Size Selection



Shell size depends on kinetic and equilibrium effects

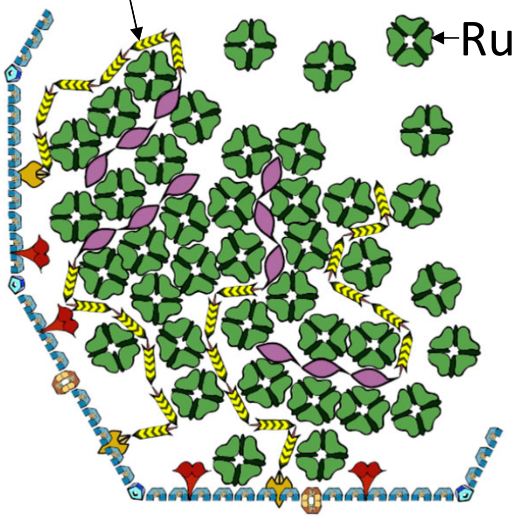
Shell-Shell affinity, ε_{SS}=1.8

Mohajerani & MFH, Plos. Comp. Biol. (2018)

Cargo topology

Farri Mohajerani

Cargo-cargo interactions are mediated by flexible scaffolds
(e.g. intrinsically disordered proteins)



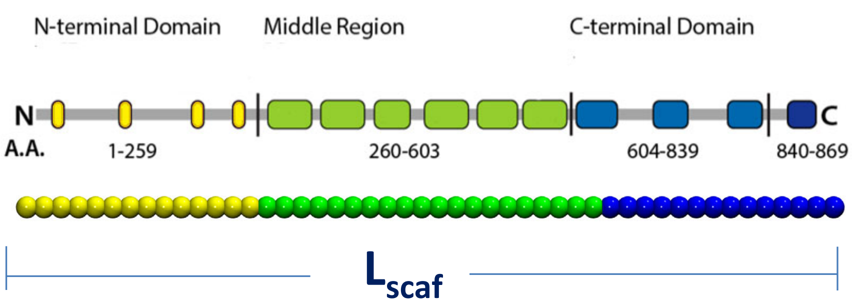
Kerfeld and Melnicki, Curr. Op. Plant Biol. 2016

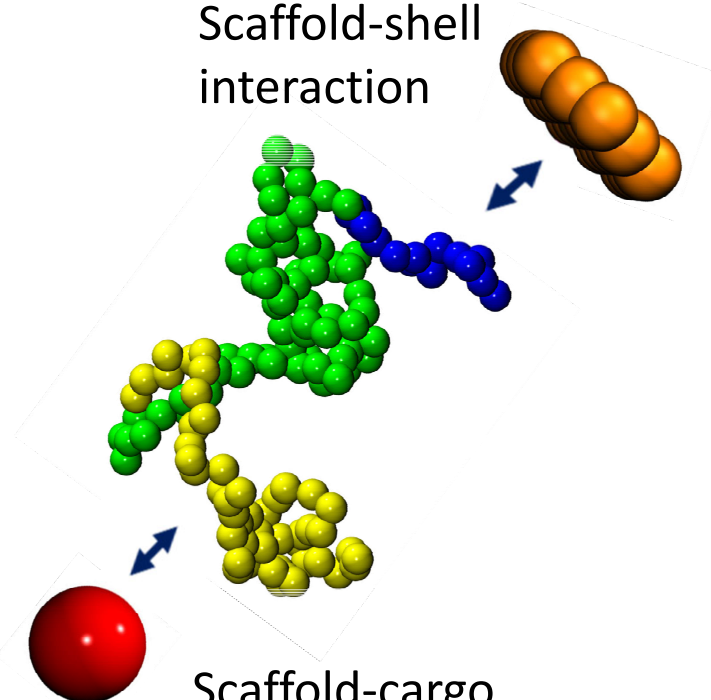
changing lengths of scaffold domains changes shell size and morphology
[L. Oltrogge & D. Savage, Nat. Struct. Mol. Biol. 2020, 27, 281−287 and unpublished]

Cargo topology

Farri Mohajerani

CsoS2, the scaffolding peptide in α-carboxysomes





Oltrogge (2020) Nat Struct Mol Biol

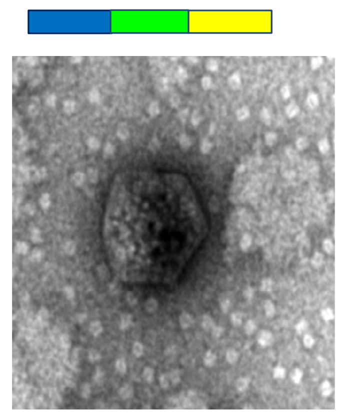
changing lengths of scaffold domains changes shell size and morphology

[L. Oltrogge & D. Savage, Nat. Struct. Mol. Biol. 2020, 27, 281–287 and unpublished]

α-carboxysome assembly with different scaffold lengths

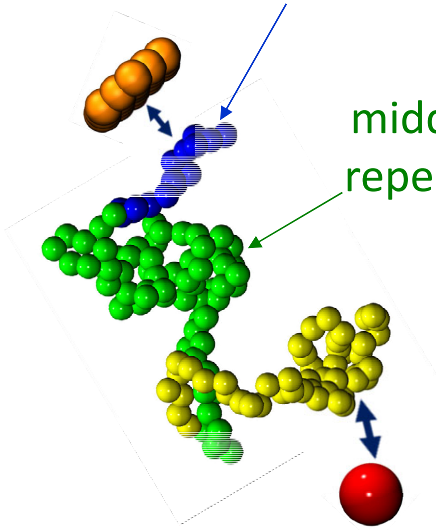
middle repeats
in scaffold

Wild-type
6



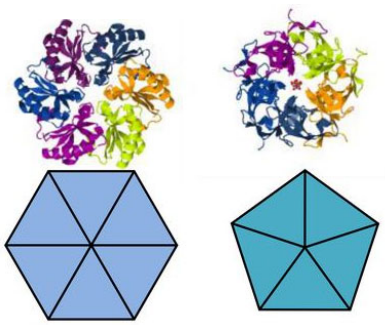
Scaffold-shell
interaction domain

middle
repeats

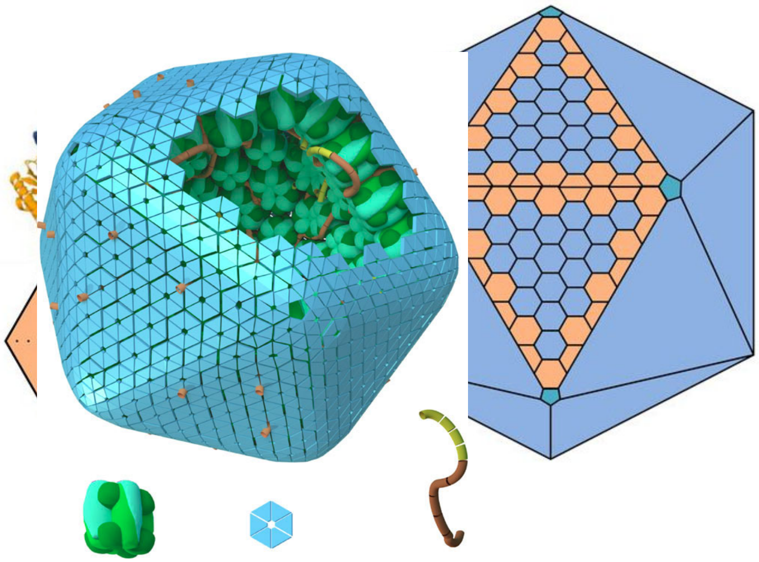


Experimental data from Luke Oltrogge, Savage Lab, UC Berkeley

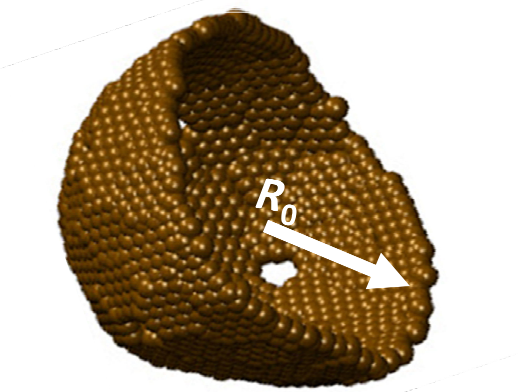
Model



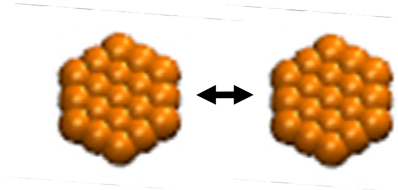
Rae (2013): *Microbiology and Molecular Biology Reviews*



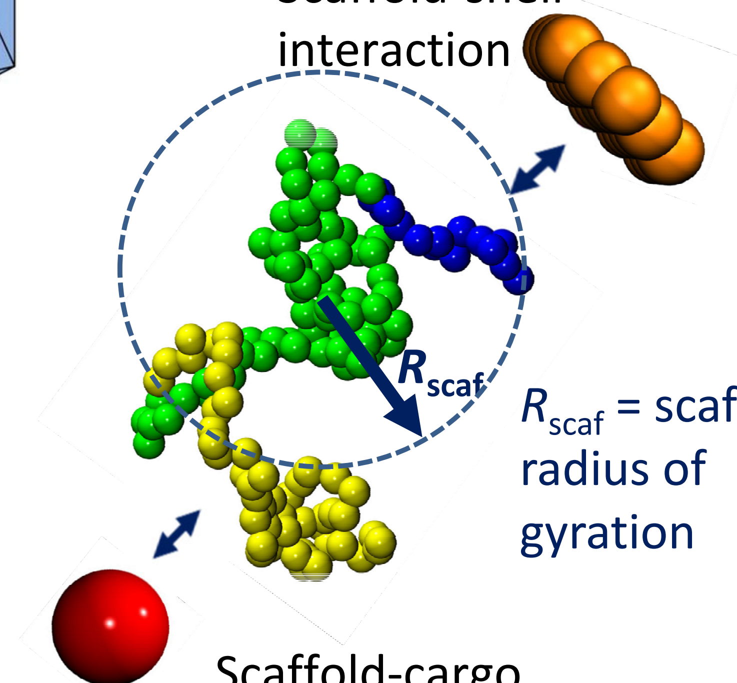
Carboxysomes can assemble from hexamers only:
Long et al. Nat. Comm. (2018)



R_0 = spontaneous radius of curvature of shell



Shell-shell interaction



Scaffold-shell interaction

R_{scaf} = scaffold radius of gyration

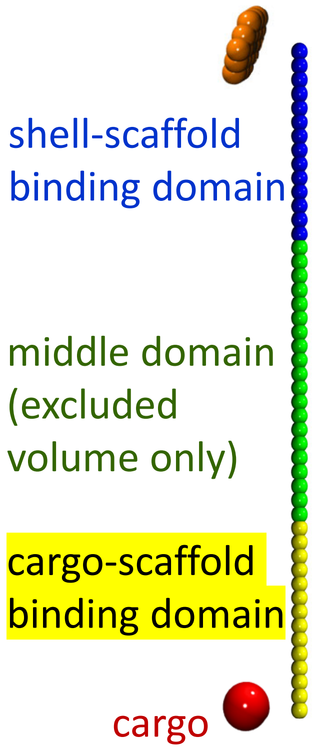
Scaffold-cargo interaction

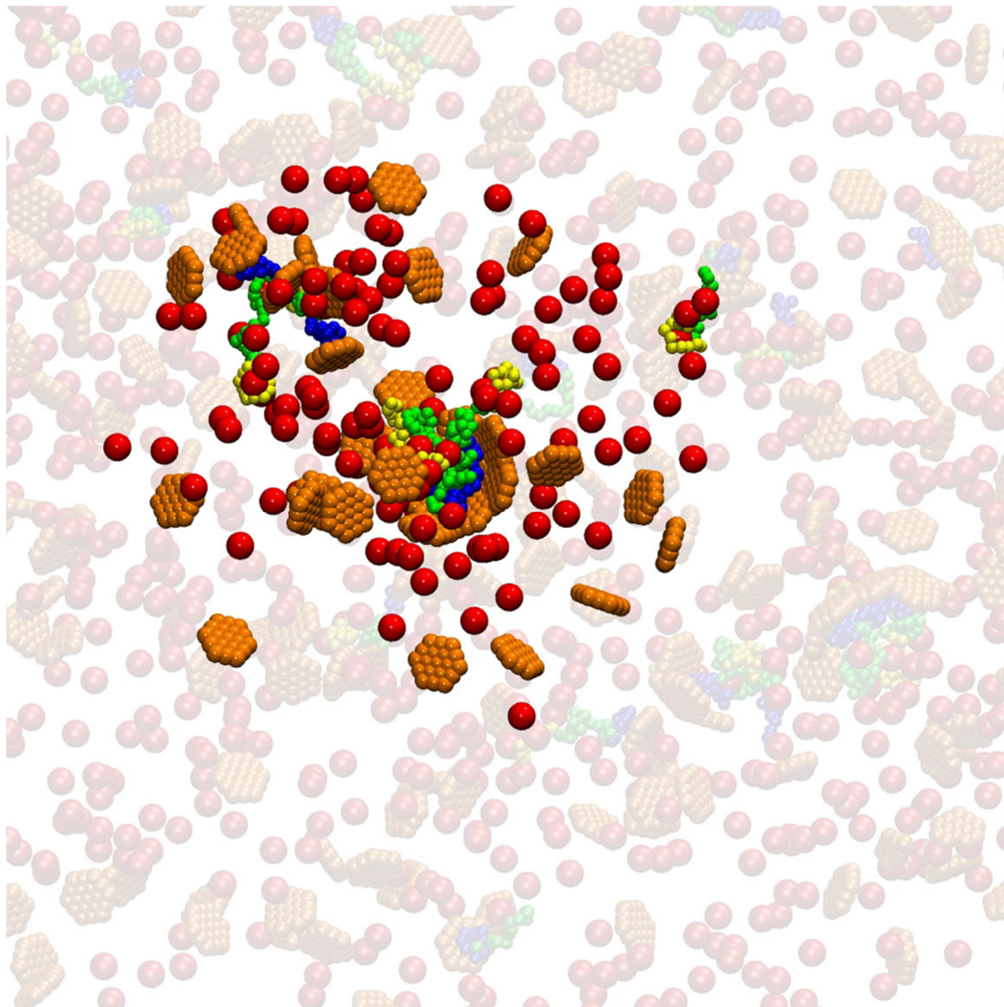
Wild-type middle domain length for α-carboxysome

$$R_0 = 8 l_{\text{rub}} \approx 100 \text{ nm}$$

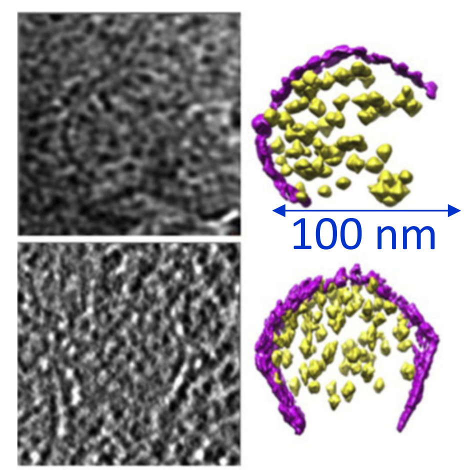
α-carboxysome diameter

$l_{\text{rub}} \approx 13$ nm is Rubisco diameter





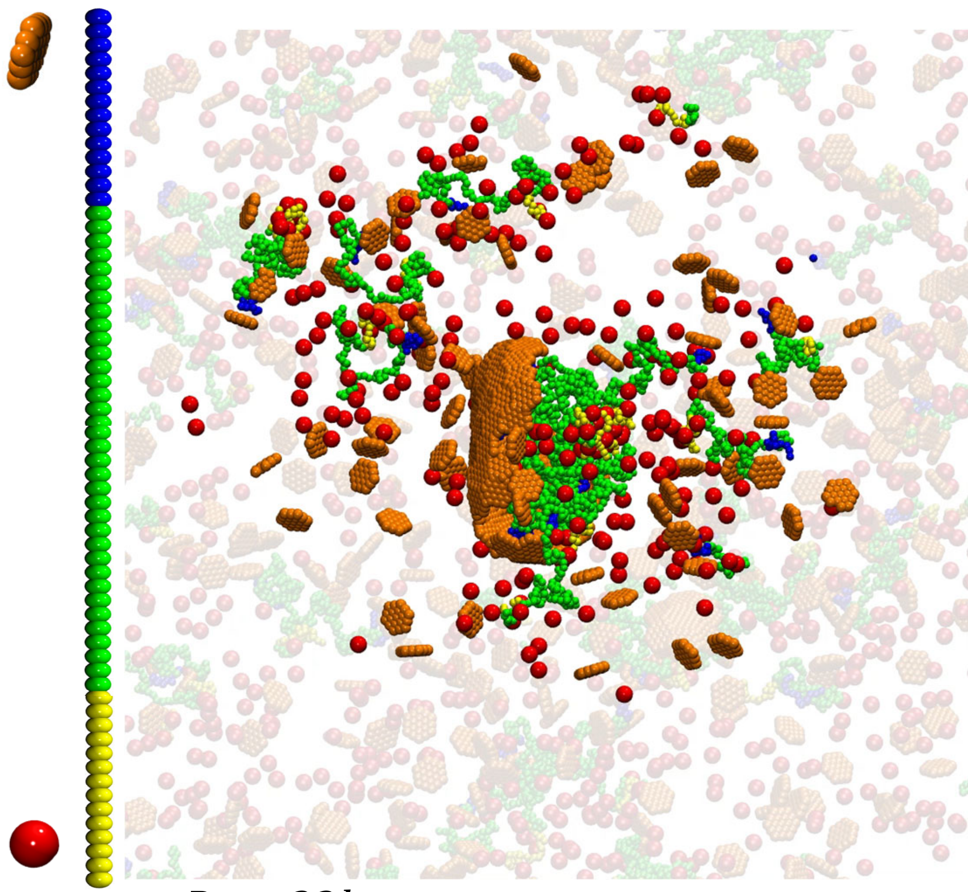
α-carboxysome assembly



Dai et al., J. Mol. Biol. (2018)
Iancu et al, J. Mol. Biol. (2010)

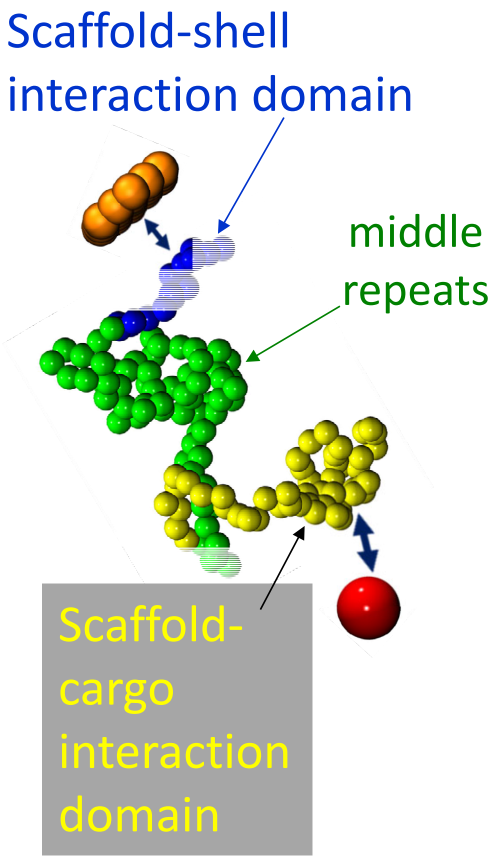
Mohajerani, Sayer, Neil, Inlow, Hagan, ACS Nano (2021)

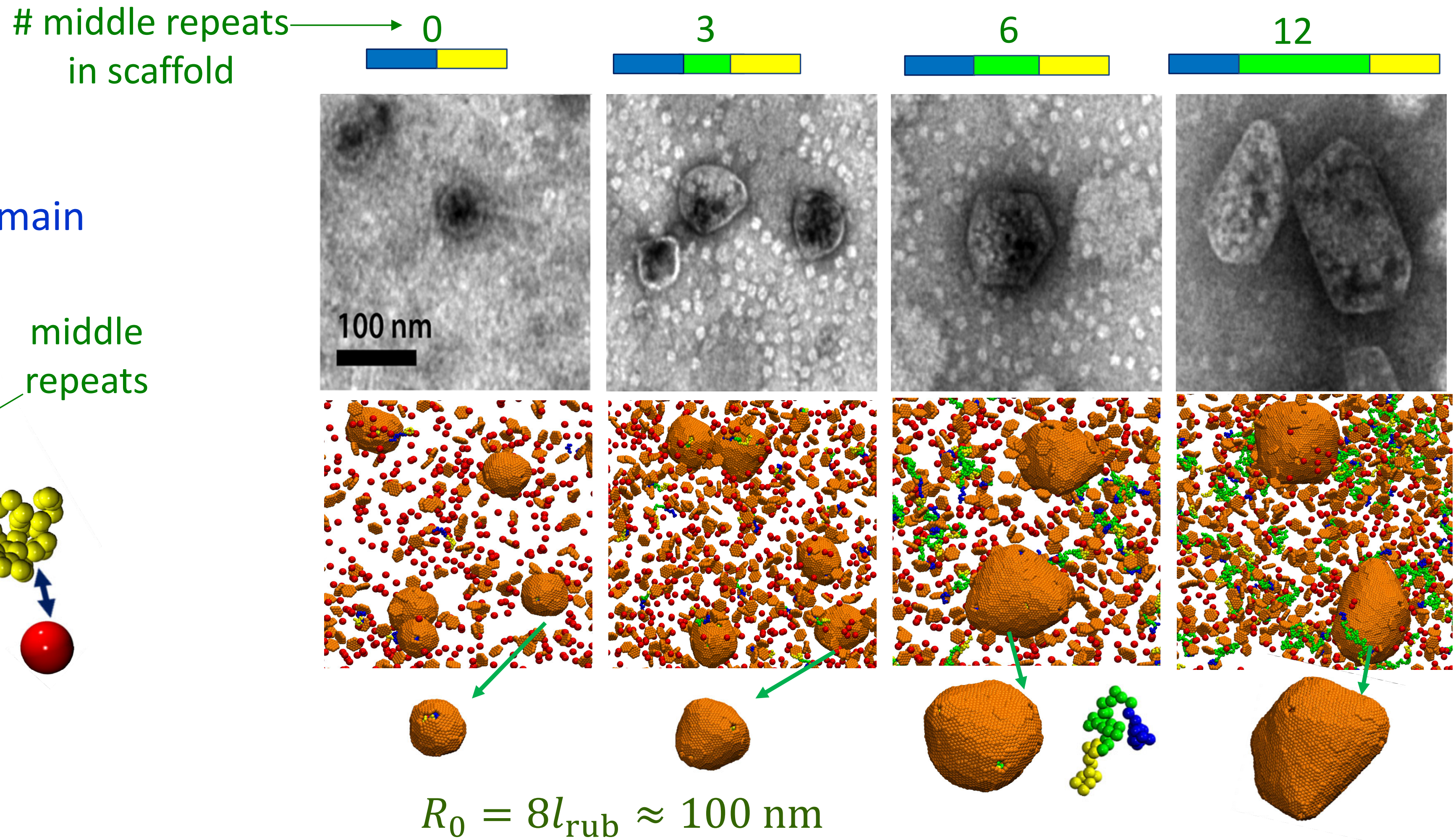
Long middle domain
-> Larger shells , low cargo loading due to excluded volume



$R_0 = 22 l_{\text{rub}}$

Carboxysome assembly with different scaffold lengths



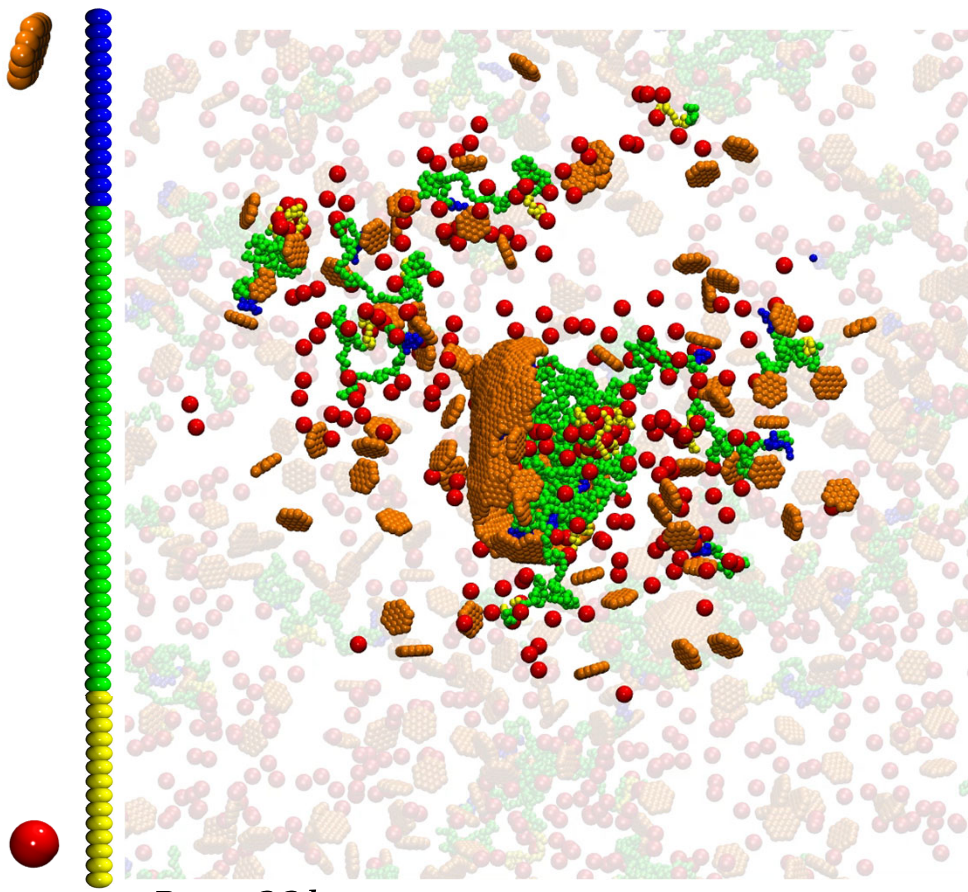


$$R_0 = 8l_{\mathrm{rub}} \approx 100\ \mathrm{nm}$$

Mohajerani, Sayer, Neil, Inlow, Hagan, ACS Nano (2021)

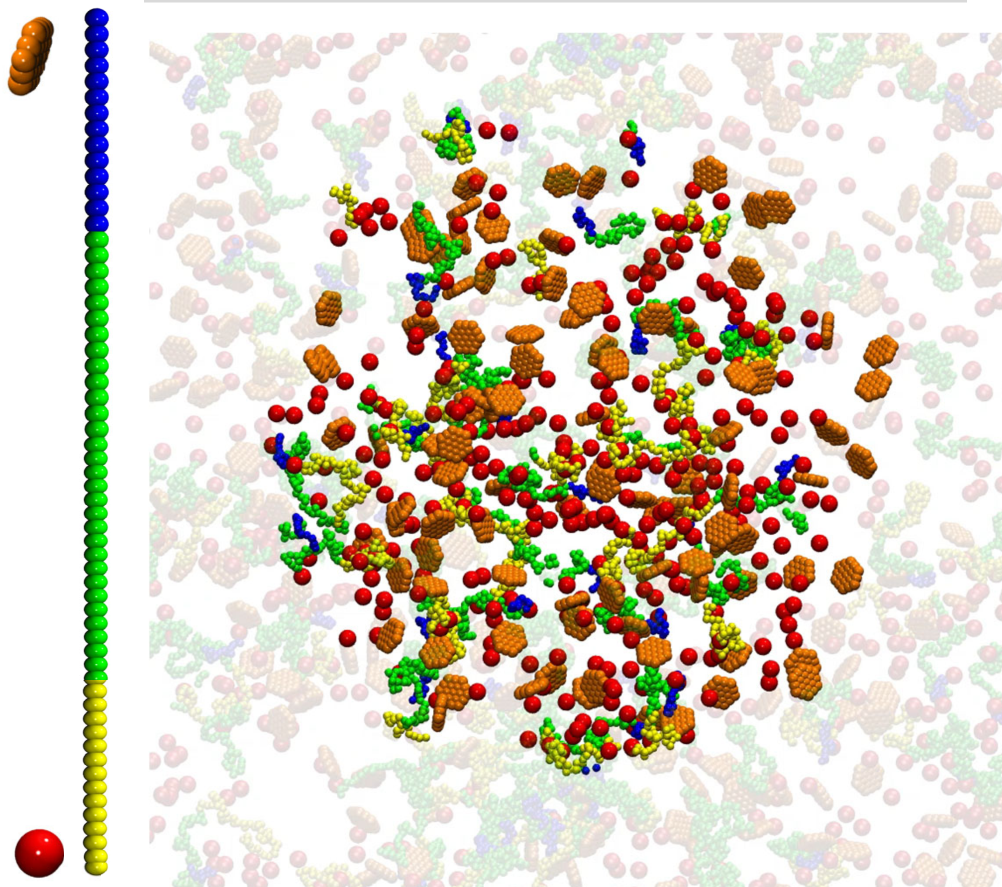
Experimental data from Luke Oltrogge, Savage Lab, UC Berkeley

Long middle domain
-> Larger shells , low cargo loading
due to excluded volume

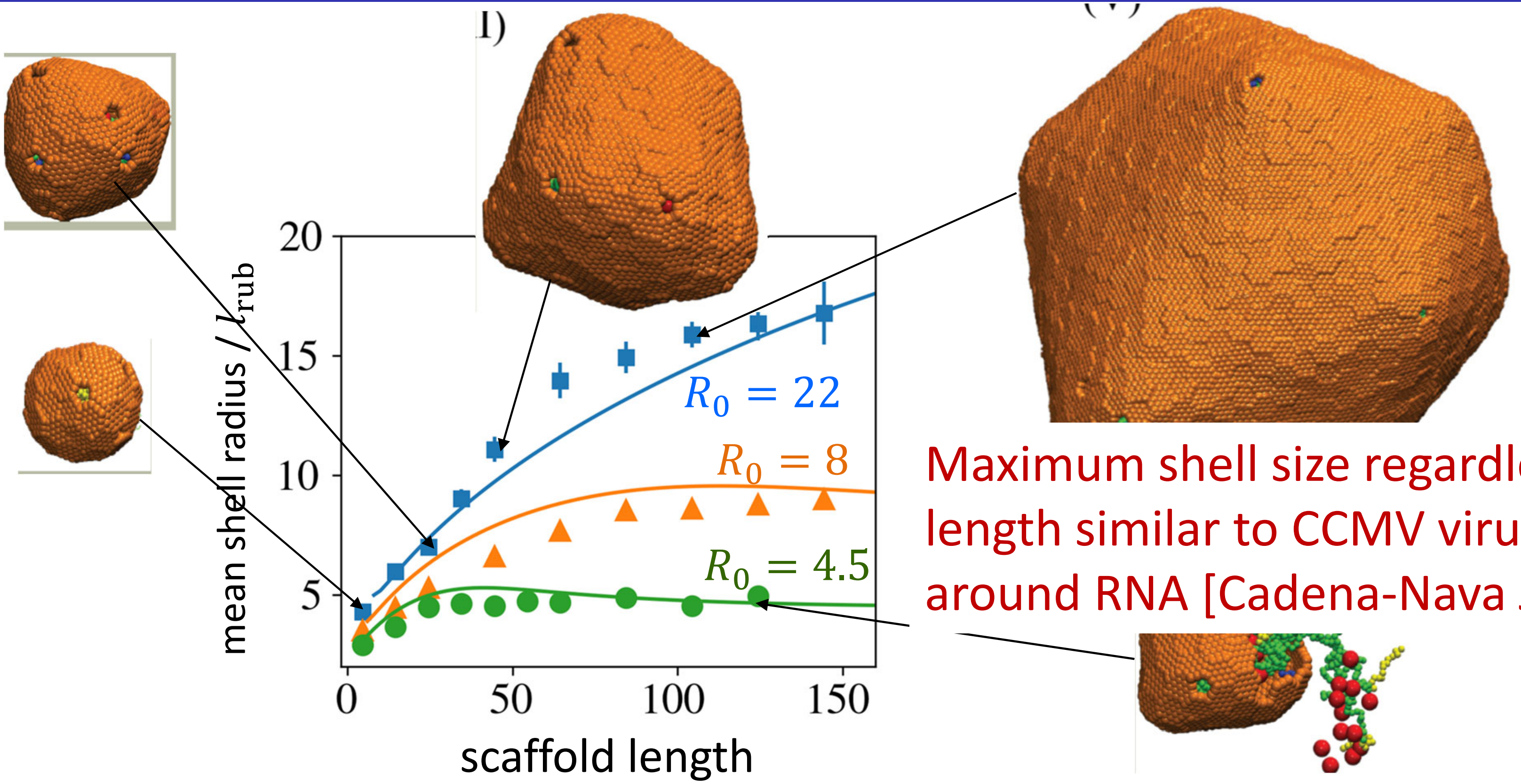


$R_0 = 22 l_{\mathrm{rub}}$

Long cargo-interacting domain
->Two-step assembly pathway
High cargo loading



Scaffold length and R_0 affect shell size



Maximum shell size regardless of scaffold length similar to CCMV virus assembly around RNA [Cadena-Nava JVI (2012)]

symbol = dynamical simulations
line = equilibrium theory

$\boldsymbol{R_0}$ =shell preferred radius of curvature

$\boldsymbol{R_0}$ sets bound on maximum shell size

Mohajerani, Sayer, Neil, Inlow, Hagan, ACS Nano (2021)

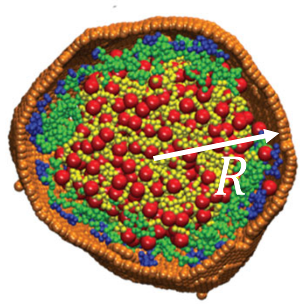
Equilibrium Shell Size Distribution

$$\rho_{n_\mathrm{H},n_\mathrm{C},n_\mathrm{S}} = \rho_\mathrm{H}^{n_\mathrm{H}} \rho_\mathrm{C}^{n_\mathrm{C}} \rho_\mathrm{S}^{n_\mathrm{S}} \exp\left[-\frac{G(n_\mathrm{H},n_\mathrm{C},n_\mathrm{S})}{k_\mathrm{B}T}\right]$$

Total free energy of assembled shell complex

concentration of shells with n_H, n_C, and n_S Hexamers, Cargo, and Scaffolds

concentrations of **unassembled** Hexamers, Cargo, and Scaffolds



$$G(n_H, n_C, n_S) = G_\mathrm{shell} + G_\mathrm{scaff} + G_{\mathrm{scaff\_cargo}}$$

$$G_\mathrm{shell} = \int dA \frac{\kappa}{2}\left(\frac{2}{R} - \frac{2}{R_0}\right)^2$$

shell bending modulus

shell radius

shell spontaneous curvature radius

$$G_\mathrm{scaff} = \left(\frac{R}{R_\mathrm{scaff}}\right)^2 + \left(\frac{R_\mathrm{scaff}}{R}\right)^2 + \ldots$$

stretching

compression

R_scaff =scaff preferred end-to-end size

Equilibrium Shell Size Distribution

shell bending modulus

$$G_{\text{shell}} = \int dA \frac{\kappa}{2}\left(\frac{2}{R} - \frac{2}{R_0}\right)^2$$

shell radius

spontaneous curvature

stretching

compression

$$G_{\text{scaff}} = n_s\left(\frac{R}{R_{\text{scaff}}}\right)^2 + \left(\frac{R_{\text{scaff}}}{R}\right)^2 + \ldots$$

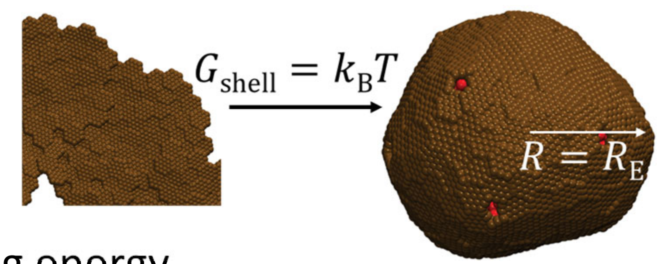
R_{scaff} =scaff size

For limiting shell subunits (ρ_{H}), minimize per-subunit energy G/n_{H}

$$\frac{g_{\text{shell}}}{k_{\text{B}}T} = \left(\frac{R}{R_{\text{E}}} - \frac{R}{R_0}\right)^2$$

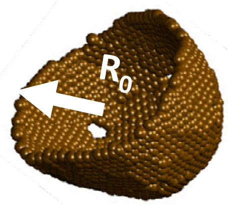
$$\frac{g_{\text{scaf}}}{k_{\text{B}}T} = \frac{n_{\text{s}}}{n_{\text{H}}}\left(\frac{R}{R_{\text{scaf}}}\right)^2 + \left(\frac{R_{\text{scaf}}}{R}\right)^2 + \ldots$$

$R_{\text{E}} \cong \sqrt{a\kappa/k_{\text{B}}T}$ a = subunit area

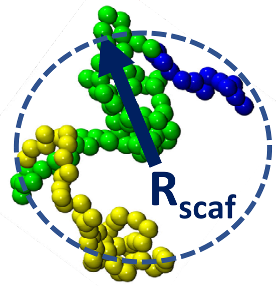


R_{E}=Elastic length: Shell size for which bending energy $G_{\text{shell}} = k_{\text{B}}T$ if no spontaneous curvature ($R_0 = \infty$)

Shell size set by competing length scales



$\boldsymbol{R_0}$ = shell spontaneous radius of curvature

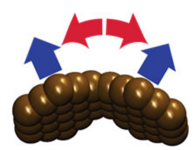


$\boldsymbol{R_{\mathrm{scaf}}}$ = scaffold radius of gyration

$\boldsymbol{R_{\mathrm{E}}} \cong \sqrt{a\kappa / k_{\mathrm{B}}T}$ Elastic length scale

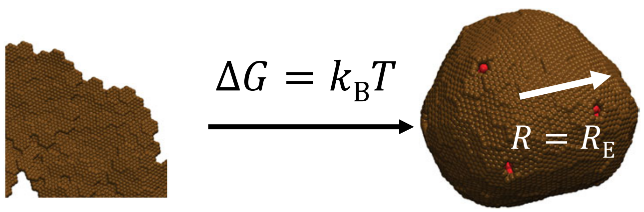
a = subunit area

κ = shell bending modulus



κ Bending modulus of shell

$\boldsymbol{R_{\mathrm{E}}}$ = shell size for which bending energy $= k_{\mathrm{B}}T$ per subunit if $R_0 = \infty$



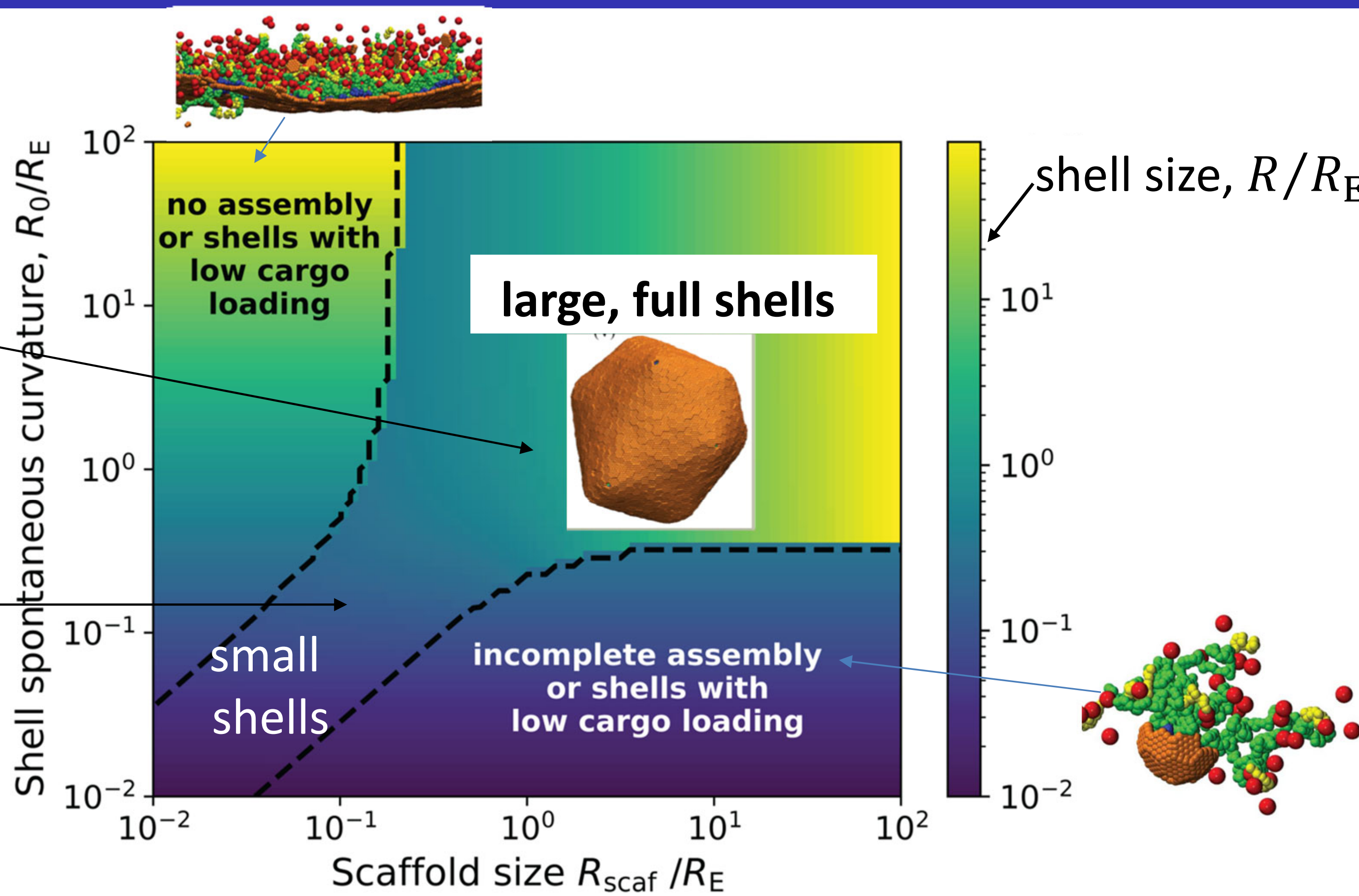
$\boldsymbol{R_{\mathrm{E}}} \approx 50$ nm for carboxysome ($\kappa \approx 25 k_{\mathrm{B}}T$, [Faulkner et al. Nanoscale 2017, 9, 10662–10673])

Mohajerani, Sayer, Neil, Inlow, Hagan, ACS Nano (2021)

Shell size set by competing length scales

$R_0 \gtrsim R_E$
shell size changes with scaffold length

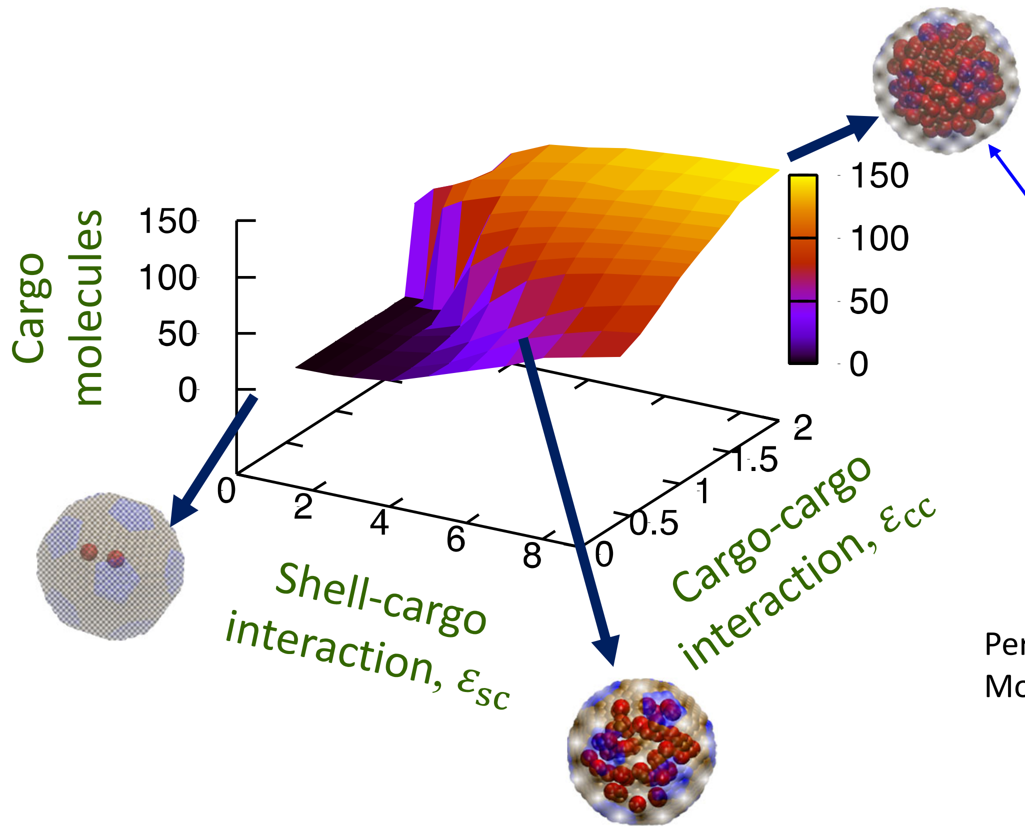
$R_0 < R_E$
shell size fixed



$R_E \cong \sqrt{a\kappa/k_BT} \approx 50$ nm is 'elastic lengthscale' of shell, depends on bending modulus

theory results set bounds on spontaneous curvature: $R_0 \gtrsim 50$ nm

Cargo Loading



High cargo loading requires cargo-cargo interactions

Perlmutter, Mohajerani, MFH, ELife (2016)
Mohajerani&MFH, Plos. Comp. Biol. (2018)

This result explains observation that targeting new enzymes to carboxysomes via shell-cargo interactions alone results in poor cargo loading

Lasilla … Kerfeld, JMB (2014)

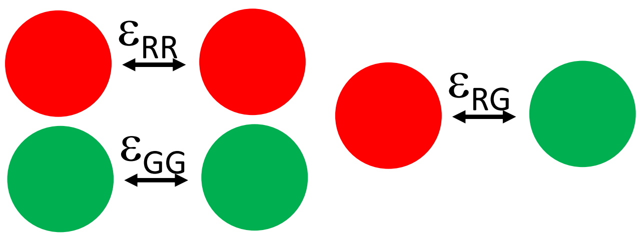
Multicomponent Cargo

Lev Tsidilkovski

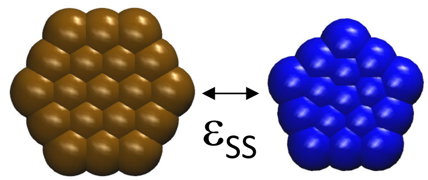
Encapsulate two cargo species, e.g. two components of reaction cascade

-Need to control their concentrations and stoichiometry!

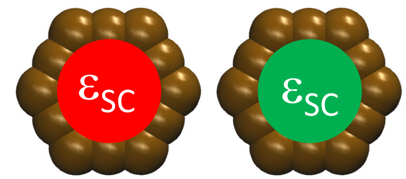
cargo-cargo interactions



shell-shell interactions



shell-cargo interactions



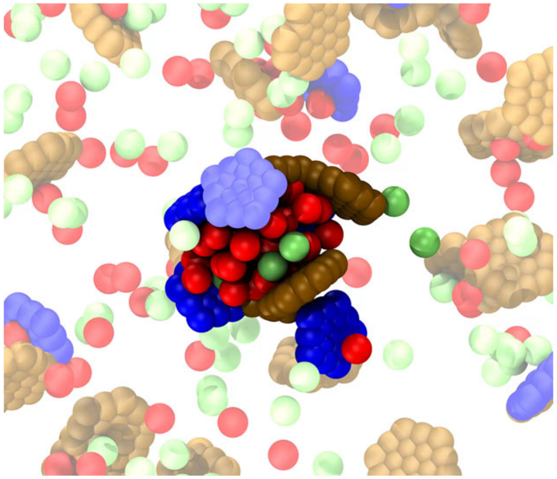
Examples and reviews of multicomponent encapsulation

- Edwardson & Hilvert, J Am Chem Soc 141, 9432 (2019).
- Patterson et al. ACS Chem. Biol. 9, 359 (2014)
- Chowdhury et al., Microbiol. Mol. Biol. Rev. 78, 438 (2014).
- C. A. Kerfeld et al., Nat. Rev. Microbiol. 16, 277 (2018).
- Slininger Lee et al., ACS Synth. Biol. (2017),
- Hagen et al., Nat. Comm. 9, 1 (2018).

Tsidilkovski, MFH, JCP 2022

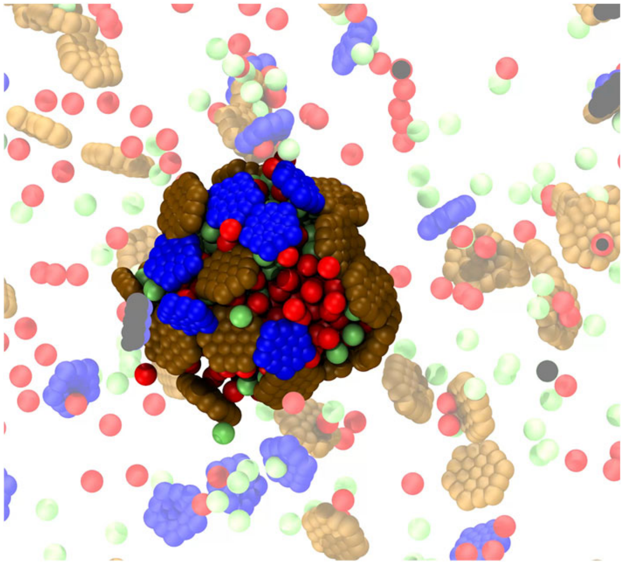
Pathways and outcomes depend on parameters

$\varepsilon_{SS} = 3.5$
$\varepsilon_{RR} = 1.7, \varepsilon_{GG} = 1.3$
$\varepsilon_{RG} = 1.3$



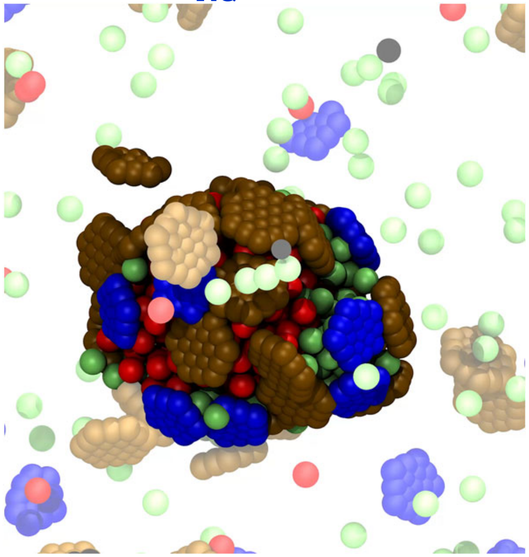
- one-step assembly
- only R cargo encapsulated

$\varepsilon_{SS} = 2.0$
$\varepsilon_{RR} = \varepsilon_{GG} = 1.7$
$\varepsilon_{RG} = 1.3$



- two-step assembly
- separation into R shells, G shells
- shell closes at interface between R/G domains

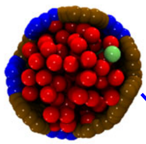
$\varepsilon_{SS} = 2.0$
$\varepsilon_{RR} = 2.0, \varepsilon_{GG} = 1.7$
$\varepsilon_{RG} = 1.3$



- two-step assembly
- R-R interactions are too strong for shell closure

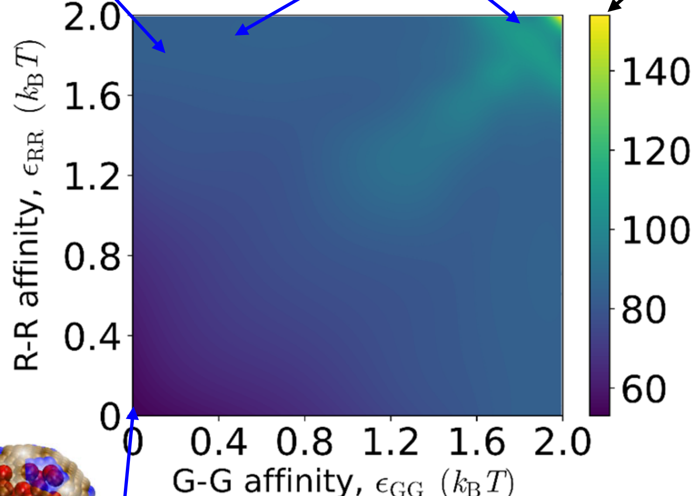
Tsidilkovski, MFH, JCP 2022

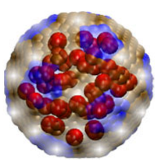
Amount of Cargo loading



High cargo loading requires strong cargo-cargo interactions

number of cargo encapsulated

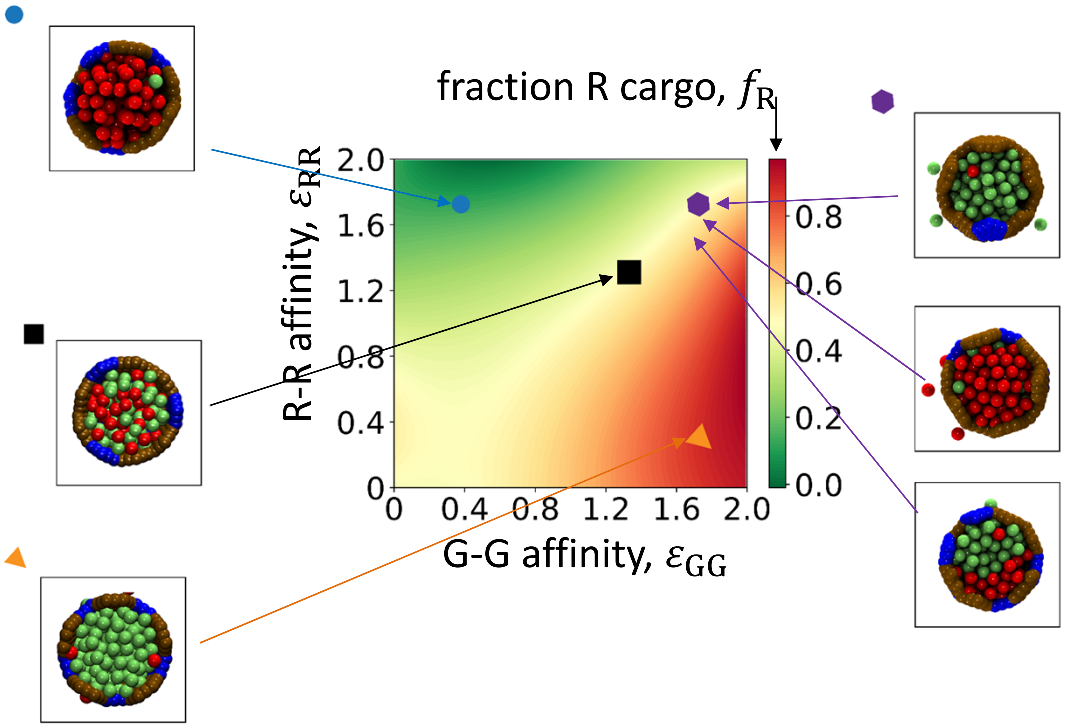


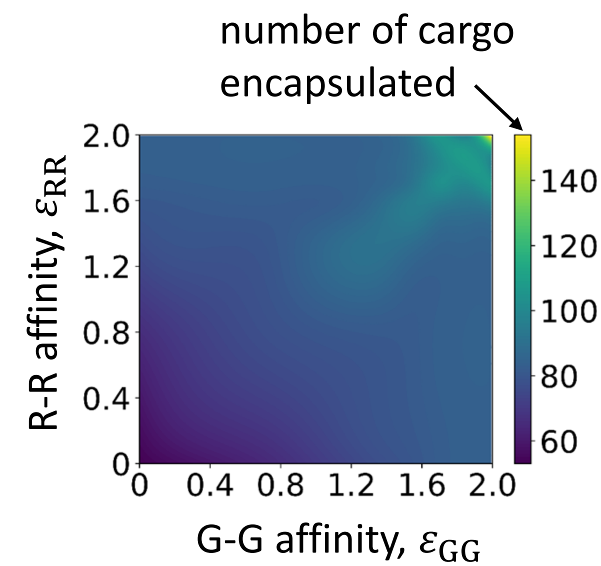


weak cargo-cargo interactions
=> only fill shell surface

This result explains observation that targeting new enzymes to carboxysomes via shell-cargo interactions alone results in poor cargo loading Lasilla … Kerfeld, JMB (2014)

Composition of encapsulated cargo





Cargo composition controlled by relative strength of $\varepsilon_{RR}, \varepsilon_{GG}, \varepsilon_{GR}$

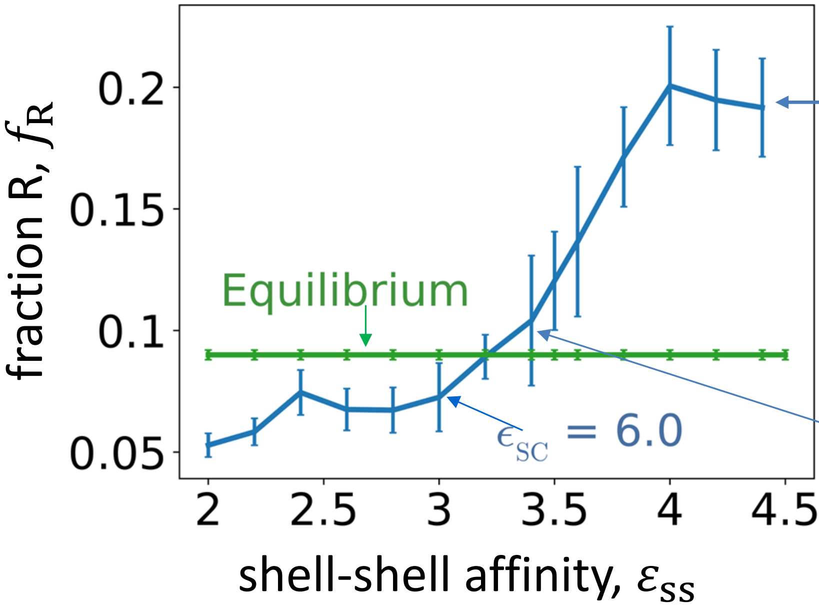
Cargo species separate into different shells when like-like interactions stronger than unlike

$\varepsilon_{RR}, \varepsilon_{GG} > \varepsilon_{GR}$

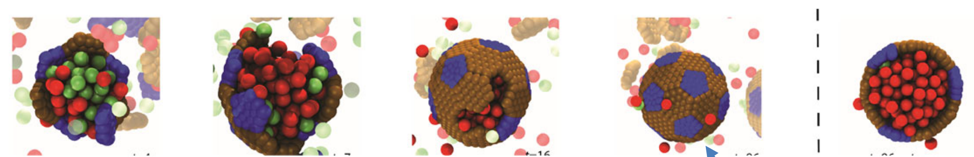
Tsidilkovski, MFH, JCP 2022

Non-equilibrium effects

both equilibrium and non-equilibrium effects control cargo composition



for strong shell-shell interactions, assembly is too fast for cargo globule to equilibrate

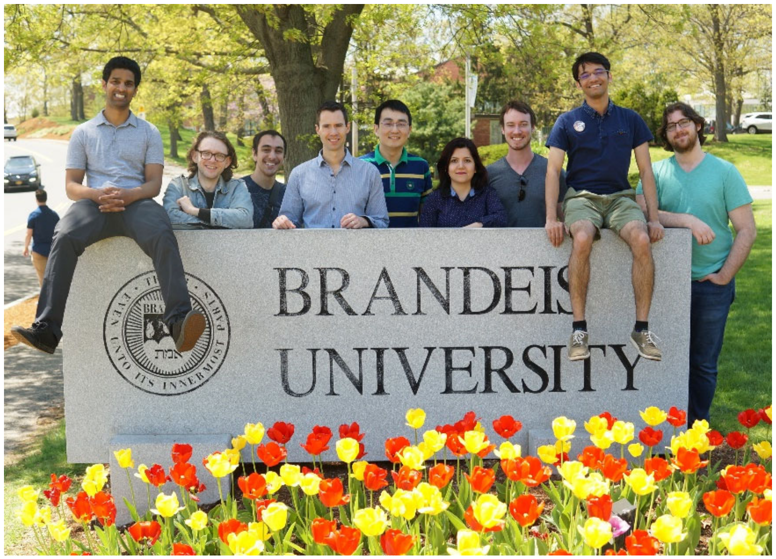


Once shell closes
cargo composition is fixed

$\varepsilon_{SS} = 3.5, \varepsilon_{RR} = 1.7, \varepsilon_{GG} = \varepsilon_{RG} = 1.3$

Acknowledgments

Hagan Group



Farri Mohajerani

Jason Perlmutter

Lev Tsidilkovski

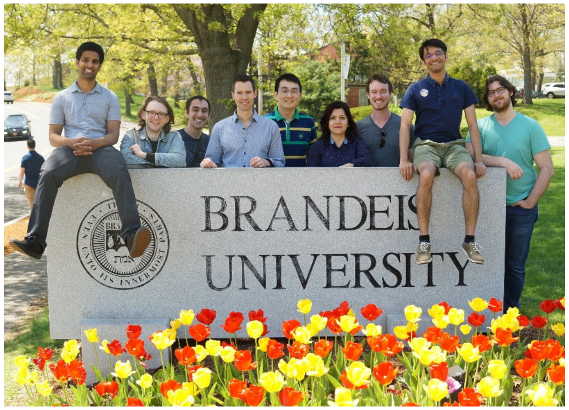
Botond Tyudoki, Evan Sayer, Chris Neil, Koe Inlow, Stefan Paquay, Anthony Trubiano

Luke Oltrogge, David Savage, Cheryl Kerfeld

$$: **NIH (R01GM108021)**
NSF (DMR-CMMT, OAC, **Brandeis MRSEC**)
DOE: Machine learning approaches to understanding and controlling 3D active matter

Computation: NSF XSEDE, Brandeis HPCC.
Simulations performed on GPUs using HOOMD

Acknowledgements



Farri Mohajerani, Lev Tsidilkovski, Jason Perlmutter, Botond Tyudoki, Evan Sayer, Chris Neil, Koe Inlow, Stefan Paquay

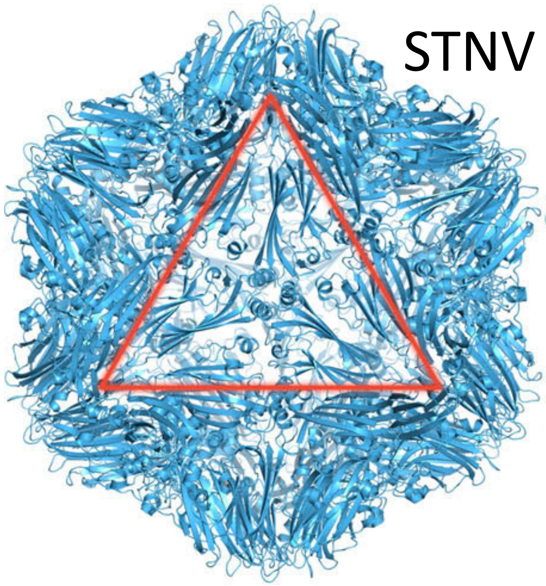
Postdoc opening: DOE project on machine learning of active matter email hagan@brandeis.edu

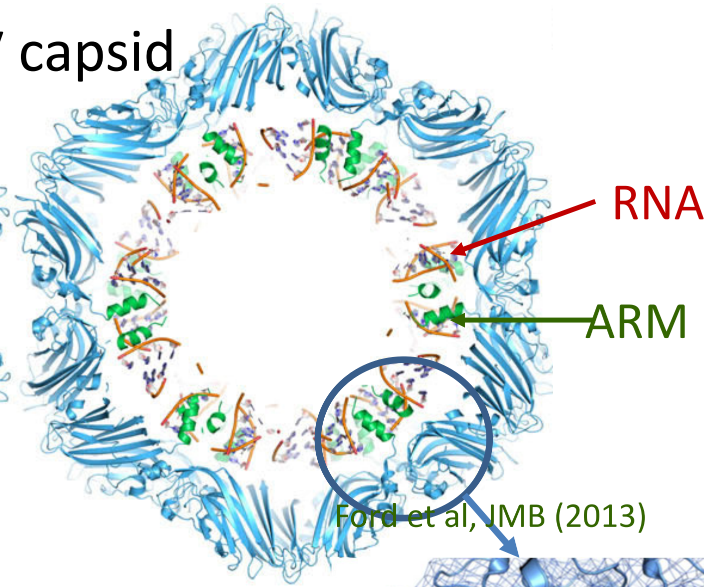
$$: NIH (R01GM108021)
NSF (DMR-CMMT, OAC, Brandeis MRSEC)
DOE: Machine learning approaches to understanding and controlling 3D active matter
Computation: NSF XSEDE, Brandeis HPCC. Simulations performed on GPUs using HOOMD

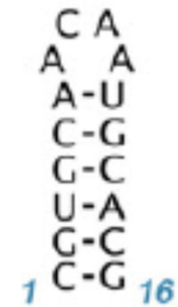
Brandeis MRSEC, AMC Crawford Notch, New Hampshire, 2020

Packaging Signals

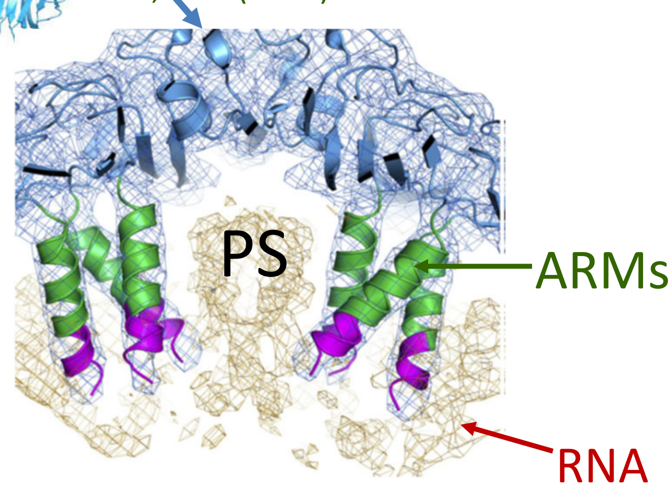
STNV capsid







Ford et al, JMB (2013)



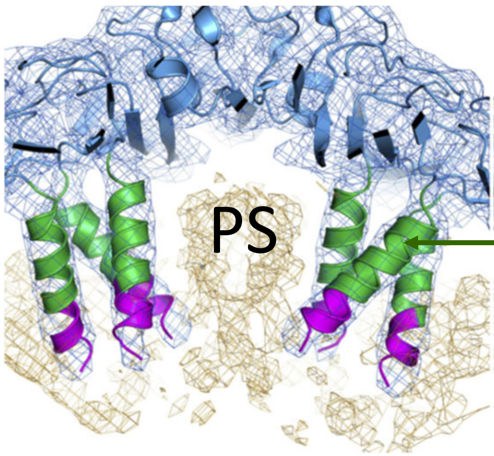
Packaging signals have two effects:

1. Bind RNA
2. Enhance protein-protein interactions

P. Stockley (Leeds),
R. Twarock, E. Dykeman (York)

Model with Packaging Signals (PS)

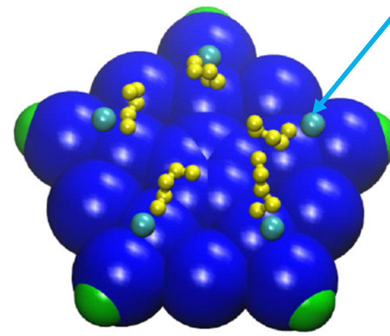
STNV Capsid with RNA



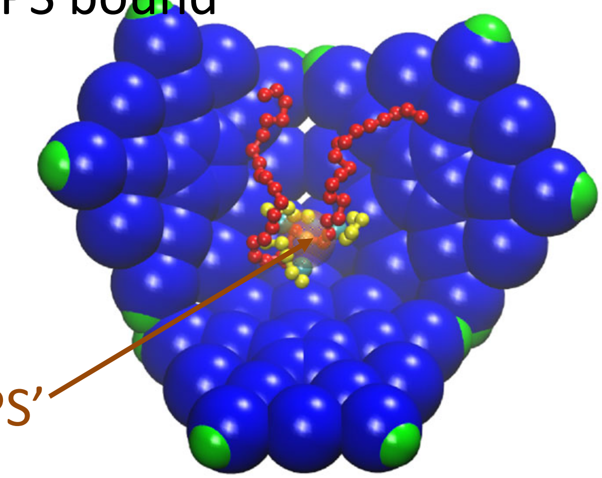
ARMs

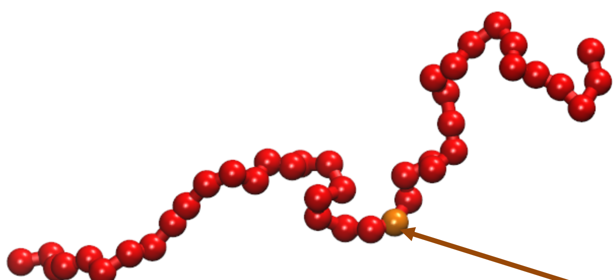
Ford et al, JMB (2013)

model PS binding sites



subunit trimer with 1 RNA PS bound

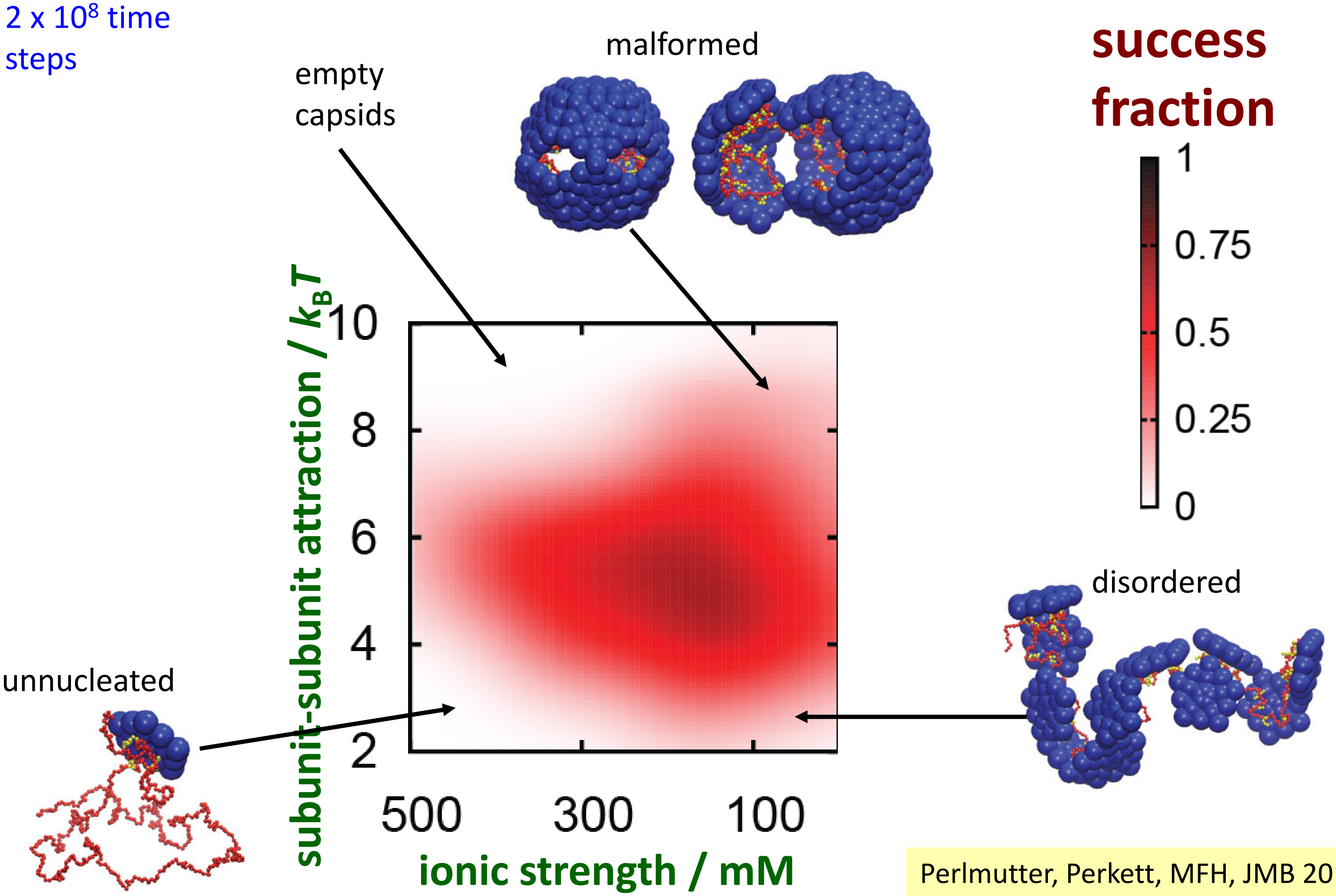




model ‘PS’

Assembly without PS

2 x 10^8 time steps

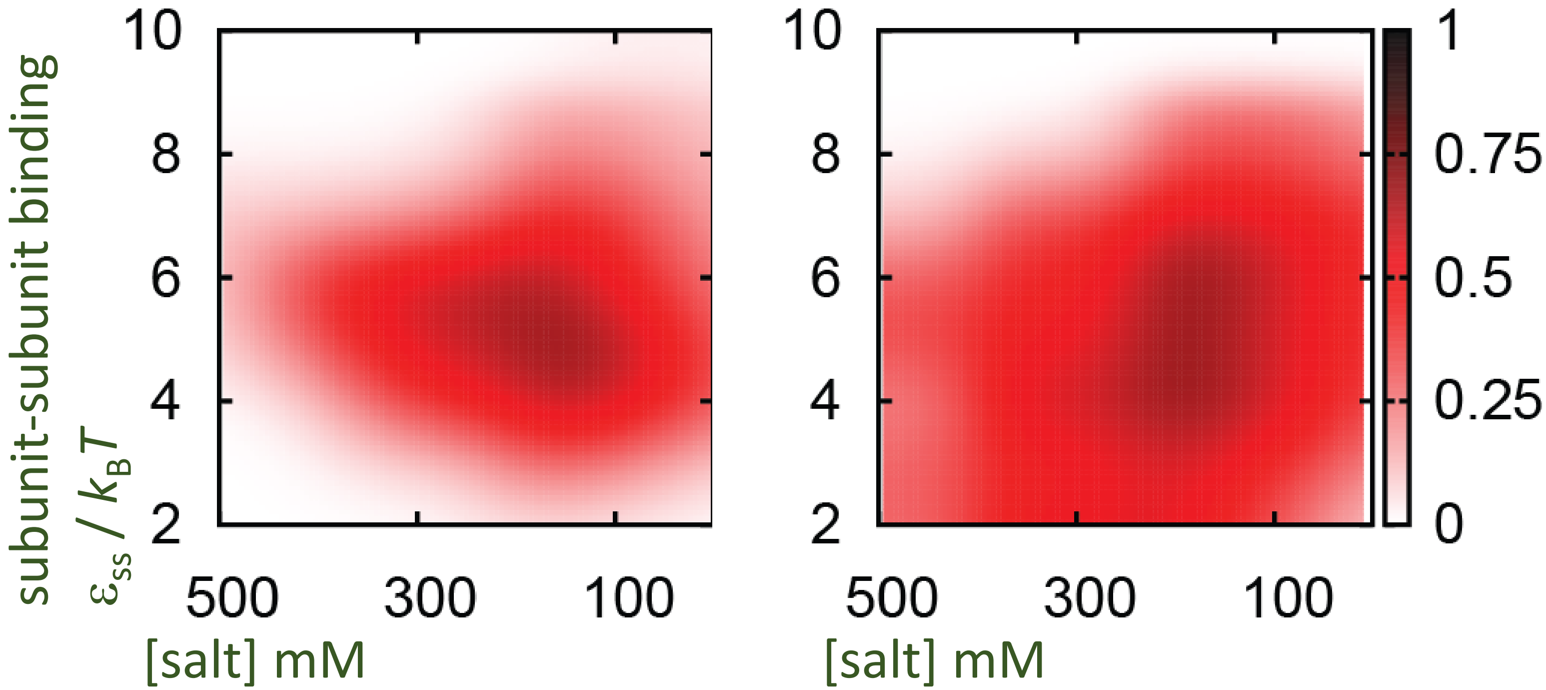


Perlmutter, Perkett, MFH, JMB 2014

Effect of PS depends on parameters

Packaging efficiency **without** PS

Packaging efficiency **with** PS

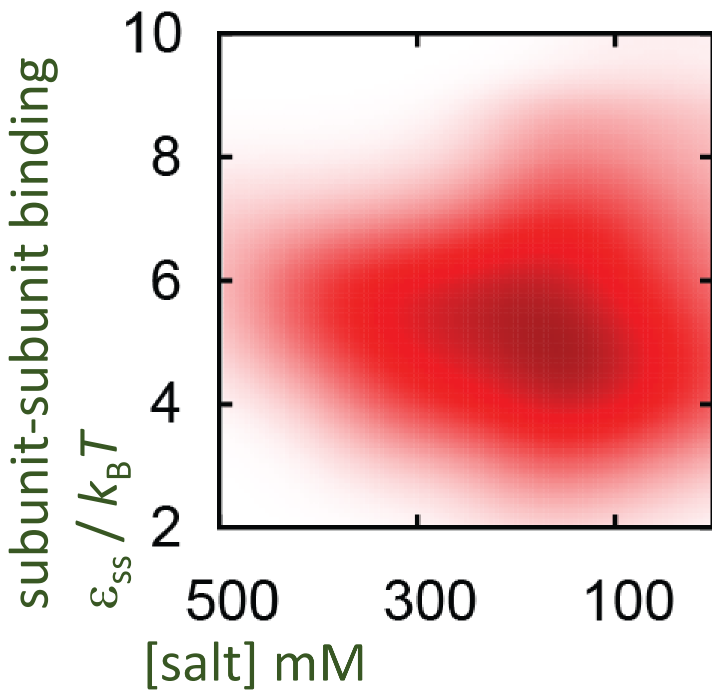


1 strong +
25 weak PS

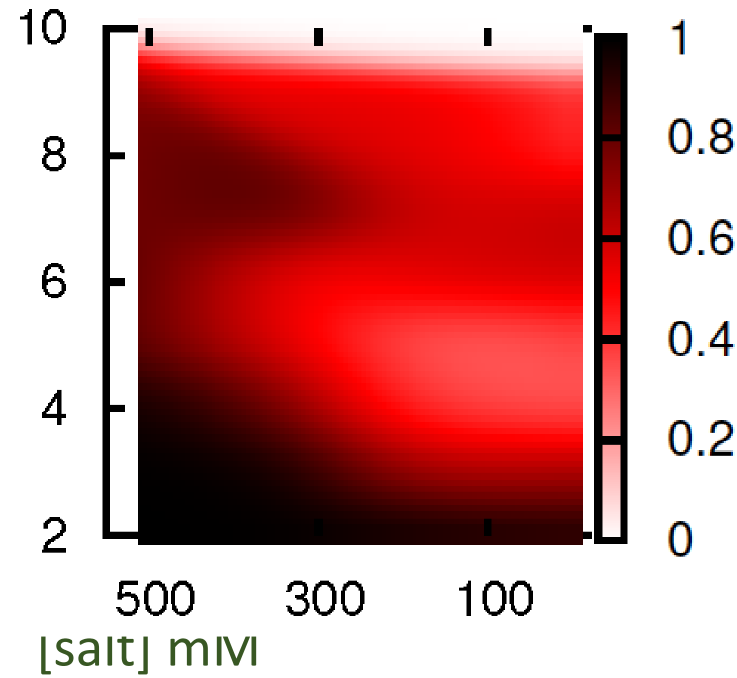
Perlmutter, MFH, JMB 2015

Effect of PS depends on parameters

Packaging efficiency **without** PS



Estimated specificity for RNA with PS



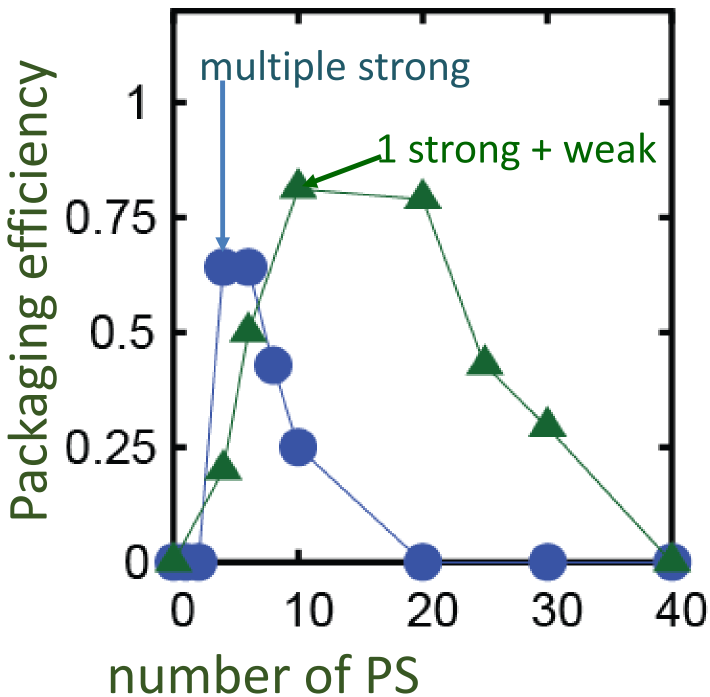
1 strong +
25 weak PS

Perlmutter, MFH, JMB 2015

Effect of PS strength and distribution

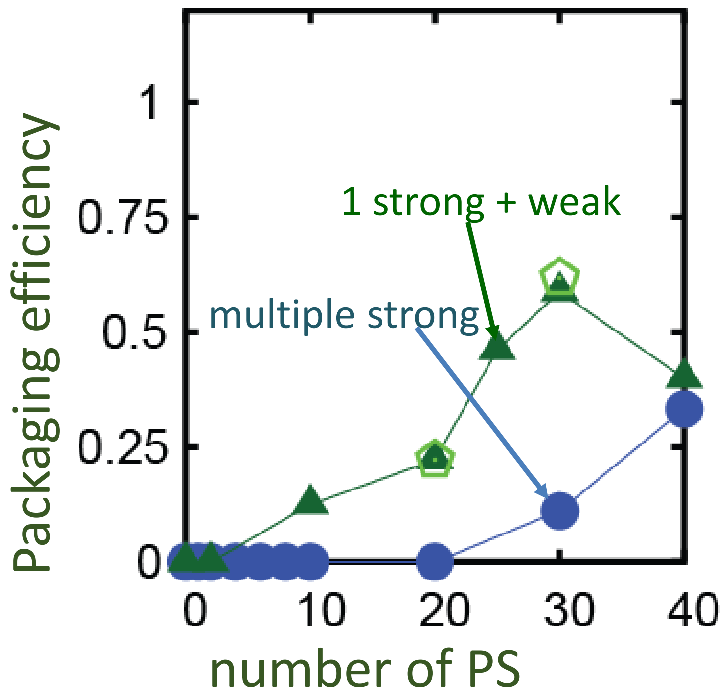
en masse mechanism

ε_{ss}=2.0 kBT, [salt]=100 mM

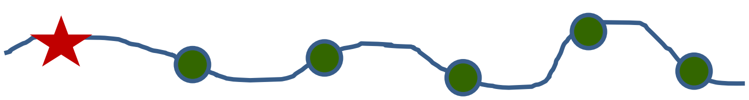


nucleation and growth

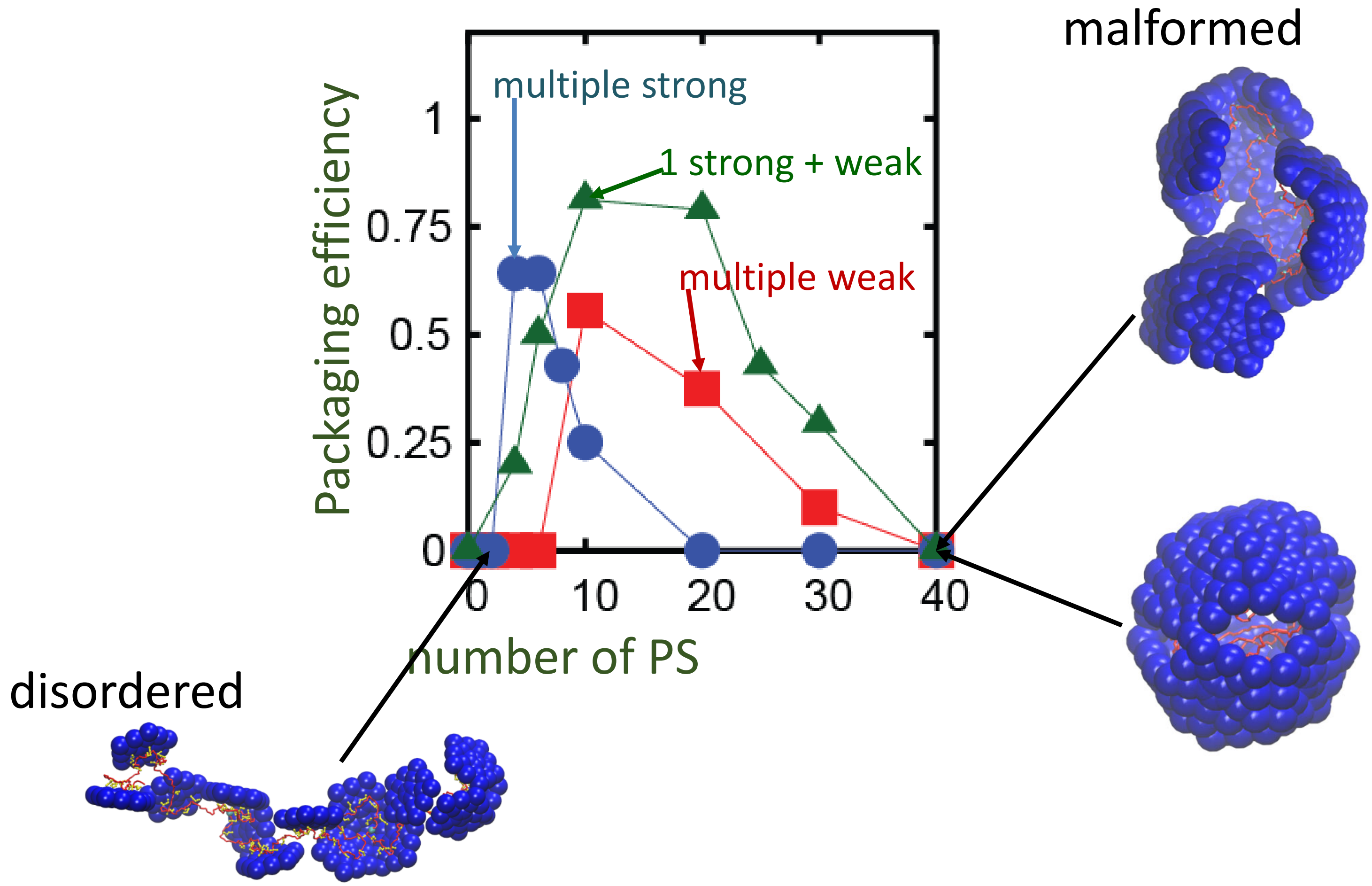
ε_{ss}=6.0 kBT, [salt]=500 mM

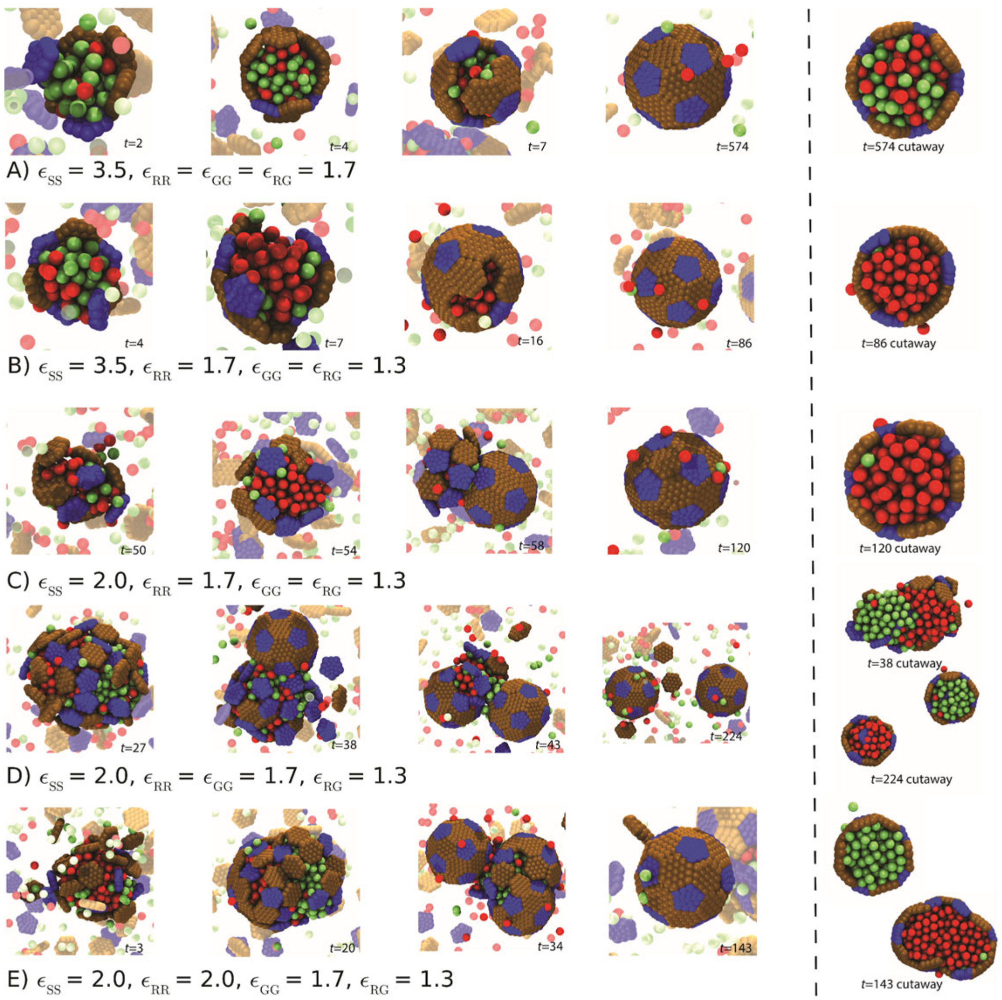


20 PS-binding sites in complete capsid



PS Failure Modes





t=2
t=4
t=7
t=574
t=574 cutaway
A) $\epsilon_{SS} = 3.5, \epsilon_{RR} = \epsilon_{GG} = \epsilon_{RG} = 1.7$
t=4
t=7
t=16
t=86
t=86 cutaway
B) $\epsilon_{SS} = 3.5, \epsilon_{RR} = 1.7, \epsilon_{GG} = \epsilon_{RG} = 1.3$
t=50
t=54
t=58
t=120
t=120 cutaway
C) $\epsilon_{SS} = 2.0, \epsilon_{RR} = 1.7, \epsilon_{GG} = \epsilon_{RG} = 1.3$
t=27
t=38
t=43
t=224
t=38 cutaway
t=224 cutaway
D) $\epsilon_{SS} = 2.0, \epsilon_{RR} = \epsilon_{GG} = 1.7, \epsilon_{RG} = 1.3$
t=3
t=20
t=34
t=143
t=143 cutaway
E) $\epsilon_{SS} = 2.0, \epsilon_{RR} = 2.0, \epsilon_{GG} = 1.7, \epsilon_{RG} = 1.3$