Computational modeling of self-assembly of Hepatitis-B Virus

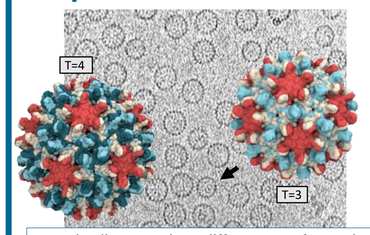
Farri Mohajerani[1] , Botond Tyukodi[1] , Christopher J. Schlicksup[2]
Jodi A. Hadden-Perilla[3], Adam Zlotnick[2], Michael F. Hagan[1]

1 Martin A. Fisher School of Physics, Brandeis University
2 Department of Molecular and Cellular Biochemistry, Indiana University
3 Department of Chemistry, University of Delaware

Assembly of the outer protein shell (capsid) of a virus is an essential step in its lifecycle. Understanding the **mechanisms underlying assembly** and the factors that determine the final morphology will guide development of antiviral drugs that disrupt or redirect assembly processes. Hepatitis-B Virus (HBV) assembles from a single capsid protein, which adopts different conformations to form icosahedral capsids with different sizes containing 180 or 240 proteins, T=3 or T=4 respectively in the Caspar-Klug nomenclature. Despite intensive experimental and theoretical investigation, the assembly pathways and mechanisms that control **HBV dimorphism** remain unclear.

I describe dynamical computer simulations of HBV assembly, using a minimal and computationally tractable model that has parameters learned from atomistic simulation data of whole HBV capsids. The simulation results identify pathways leading to T=3 and T=4 capsid and other morphologies, and suggest key factors which control this dimorphism. **Our approach is generalizable** to other self-assembling systems and may guide the engineering efforts for the **optimal design** of subunits resulting in the **robust assemblv of nanoshells**.

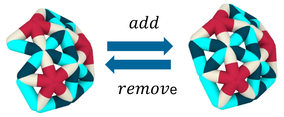
Hepatitis-B Virus



Protein dimers adopt different conformations and assemble into capsids

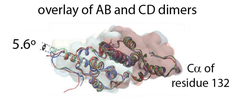
T=4, more numerous, infectious
T=3, less numerous, smaller[1]

Monte-Carlo Simulations



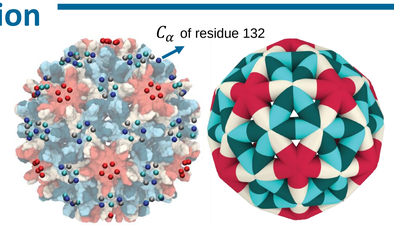
Grand canonical Monte-Carlo simulations with reversible dynamics[2]

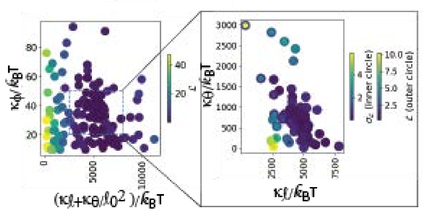
overlay of AB and CD dimers



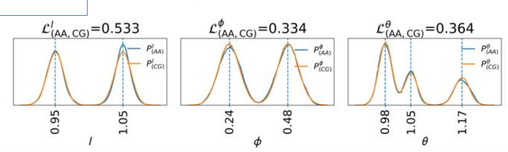
Parameter optimization

To optimize the model parameters, we perform simulations on a closed shell with a fixed topology corresponding to a T=4 icosahedral capsid. After equilibrating the shell, we optimize parameters by minimizing the difference between the distributions of thermal fluctuations in the CG model simulations and the AA molecular dynamics trajectories[3] on a T=4 HBV structure.



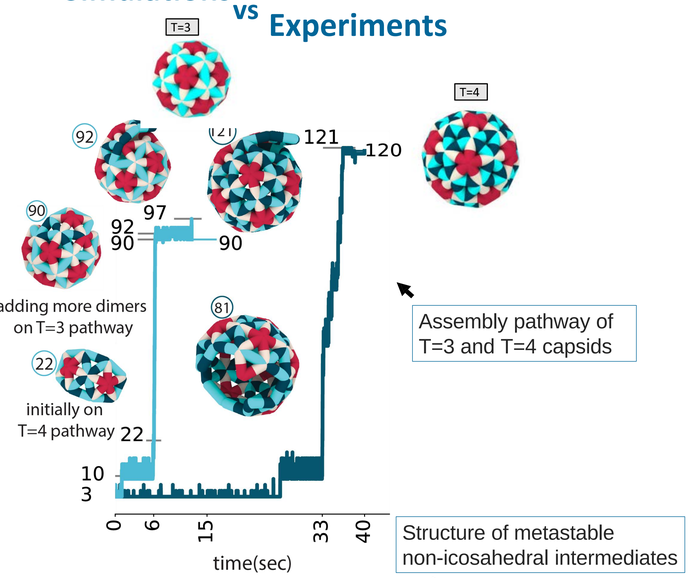


$\kappa_l = 4200, \kappa_\phi = 40, \kappa_\theta = 800, \mathcal{L}_{(AA,CG)}=0.410$



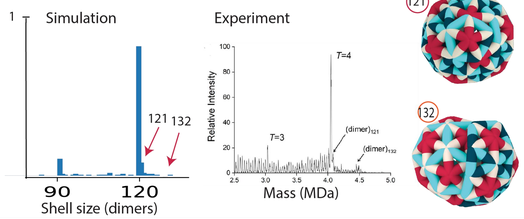
IRG1 : Optimize the model parameters from the distribution of shell sizes in experiments

Simulations vs Experiments

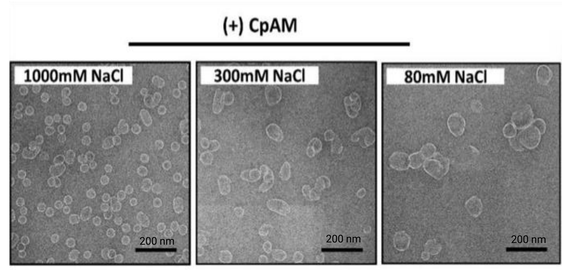


Assembly pathway of T=3 and T=4 capsids

Structure of metastable non-icosahedral intermediates



Antiviral Agents: Core protein Assembly modulators (CpAms)

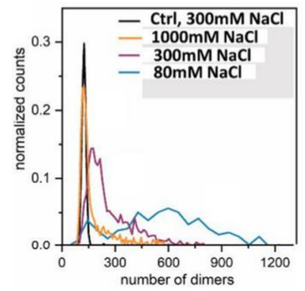


Strong assembly conditions:
- Smaller products
- Higher curvature
- (largely T=4 wildtype)

Proposed Mechanism: Speed up assembly towards aberrant non-infectious structures.

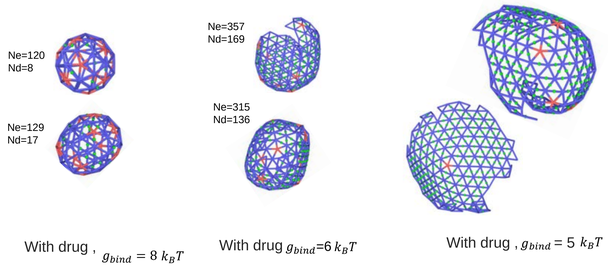
Weak assembly conditions:
- Larger products
- lower curvature[4]

Product distribution by single molecule nanopore sensing [1]



Simulations of shell assembly with drugs that bind edges and change the preferred angle of the edges result in aberrant structures. Larger structures are observed by lowering the binding affinity of shell dimers (g_{bind}) .

Ne: number of edges in the shell, Nd: number of drugs in the shell

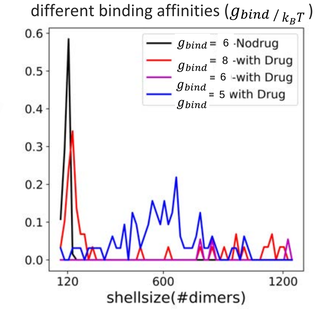


With drug , $g_{bind} = 8\ k_BT$

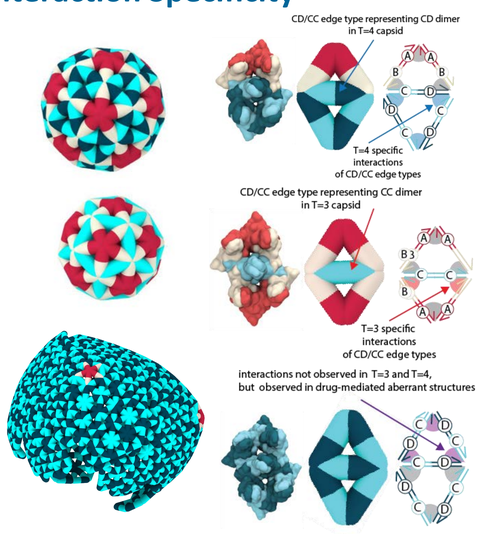
With drug g_{bind}=6 k_BT

With drug , g_{bind} = 5 k_BT

Distribution of simulation outcomes at different binding affinities (g_{bind} / k_BT)

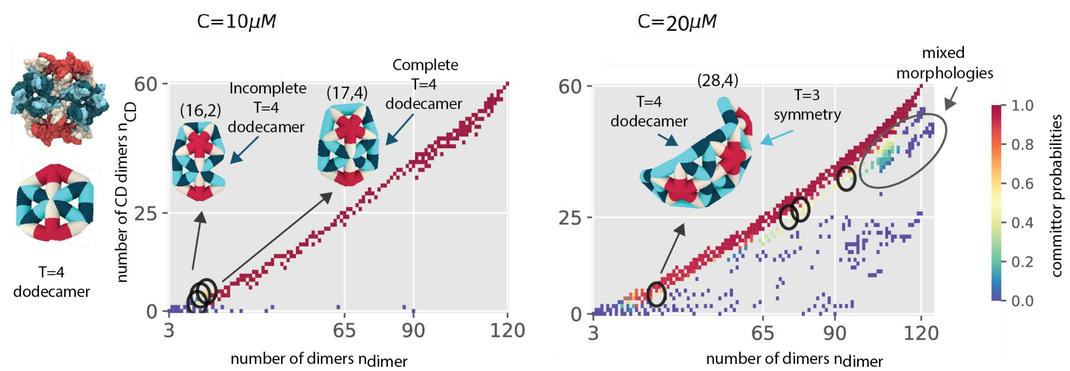


Interaction Specificity



IRG1 : Optimal subunit design and high assembly yields can be achieved by interaction specificity.

Markov State Models For Pathway Analysis



- Commitor probabilities identify pathway selection hub states.
- The majority of states close to T=4, end up in T=4 at the lower concentration, but at higher dimer concentrations, a fraction of these states remain trapped in other morphologies.
- Formation of a `T=4 dodecamer is the key event that determines whether pathways are more likely to proceed to T=4 or T=3 products.[5]

IRG1 : Markov State Models and Transition Path Theory can reveal critical information about the assembly pathways and guide the self-assembling systems toward the target structures.

[1] Zlotnick, A. *et al. Biochemistry* 1996. [2] Rotskoff, G.M. *et al.* PNAS 2018. [3] Hadden, J.A. *et al.* eLife 2018. [4] Kondylis, P. *et al.* JACS 2018. [5] Mohajerani, F. *et al.* bioRxiv 2022.

We acknowledge NIH award number R01GM108021, Brandeis MRSEC, DMR-2011846, NSF XSEDE computing resources, and the Brandeis HPCC.