DNA in Tight Spaces: From Nucleosome and Chromosomes to Origami and Viruses

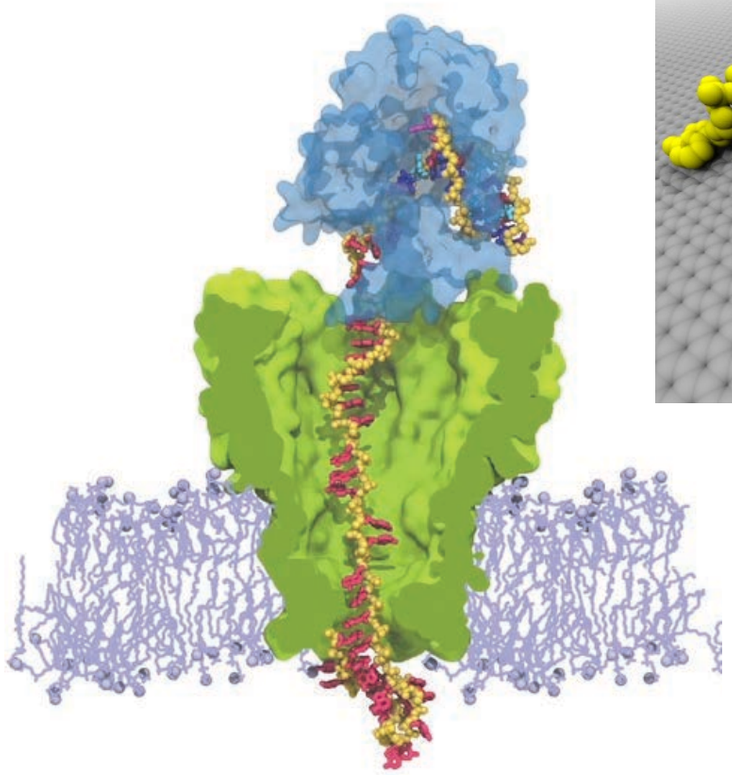
Aleksei Aksimentiev
University of Illinois at Urbana-Champaign

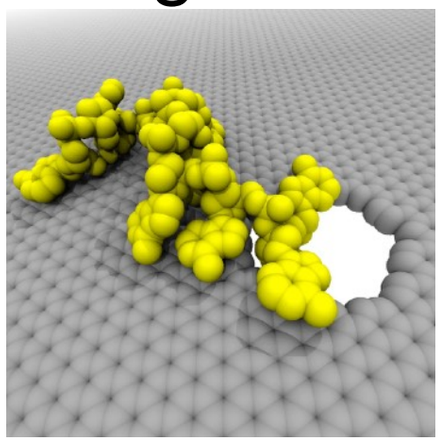
Center for the Physics of Living Cells

NSF

Other research areas

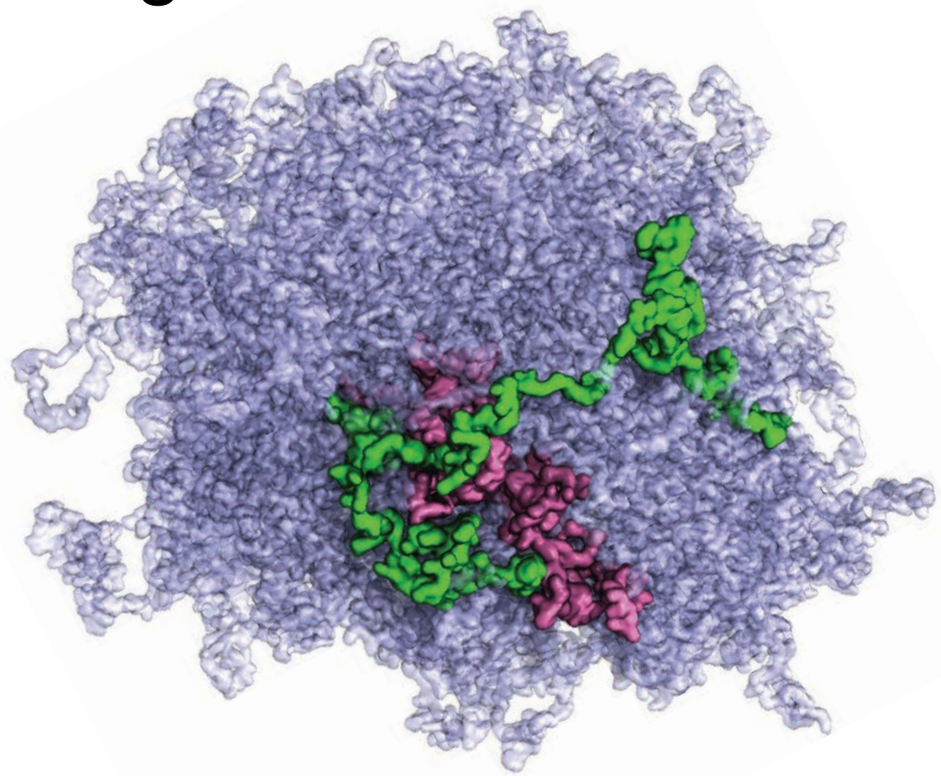
Nanopores sequencing



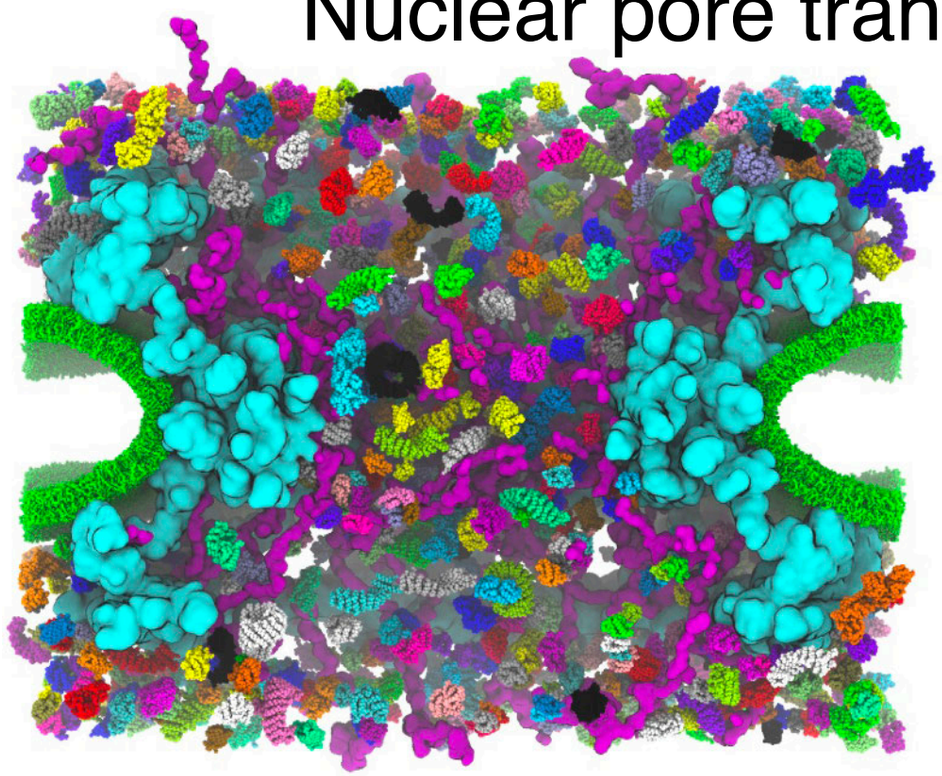


minION

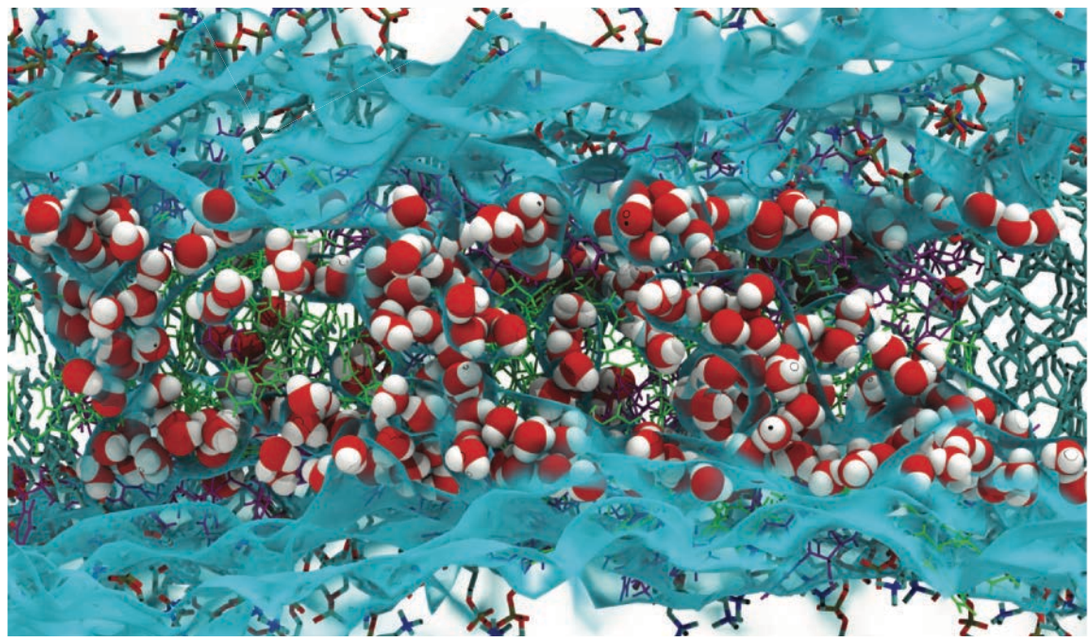
Biological condensates



Nuclear pore transport



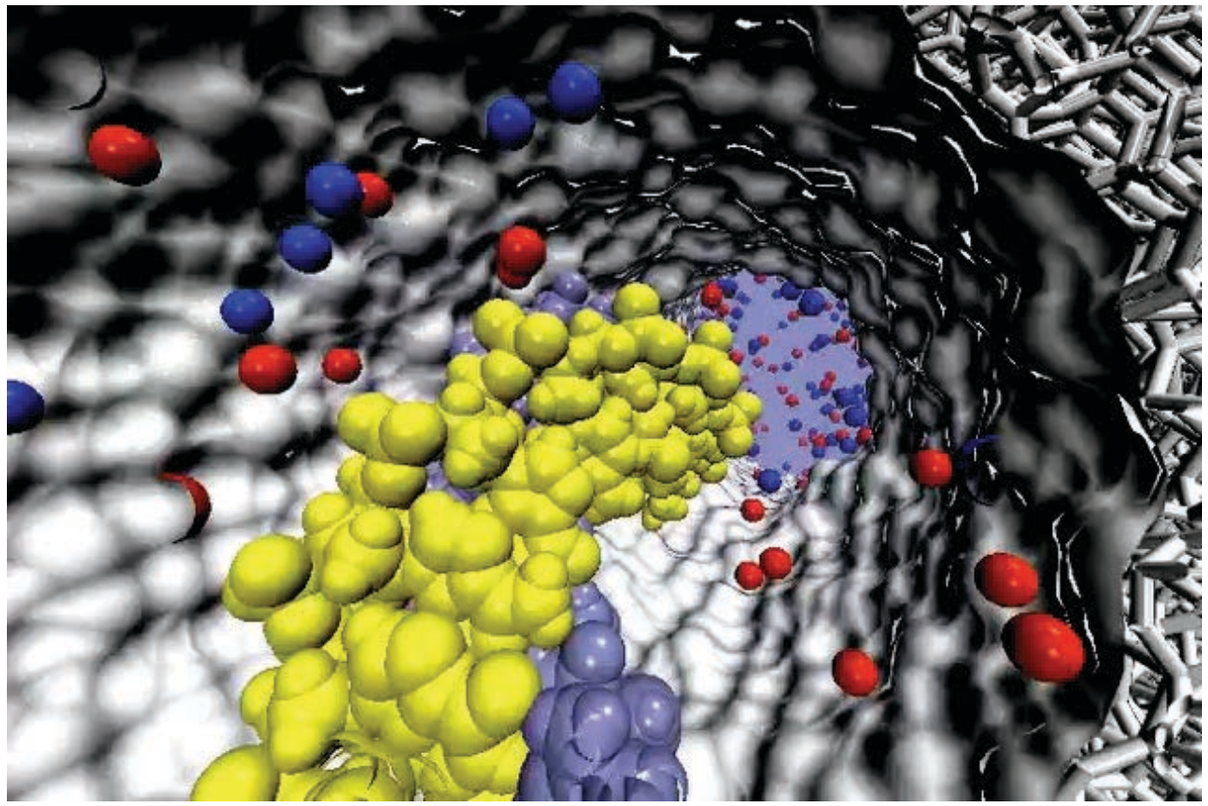
Biomimetic membrane channels



All-atom molecular dynamics simulations: the computational microscope

Massive parallel computer
Blue Waters (UIUC): ~200,000 CPUs

Atoms move according to
classical mechanics (F= ma)



Interaction between atoms is
defined by molecular force field

Time scale: ~ 0.1-100 μs

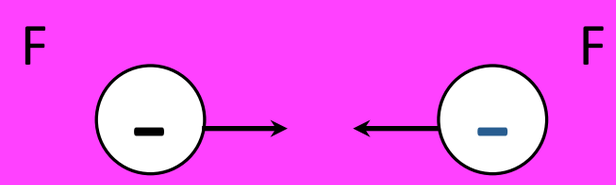
Length scale: 10K - 200M atoms or (< 100 nm)^3

Time resolution: 2 fs

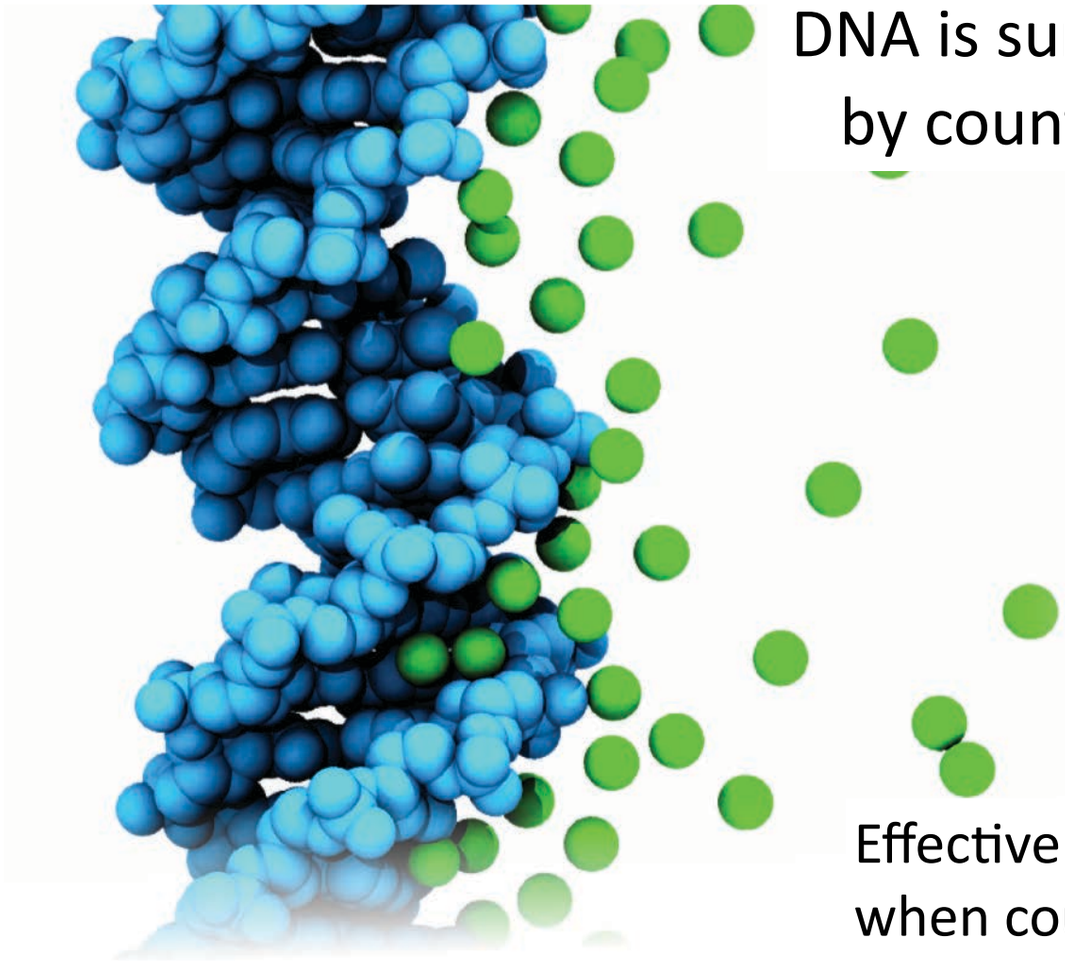
Spacial resolution: 0.1 A

Same sign charges ….

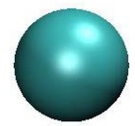
Same sign charges repel
(in vacuum)



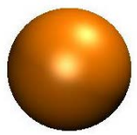
Same sign charges can
attract (in a medium)



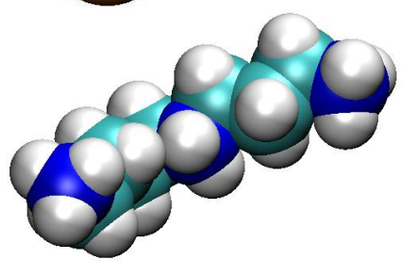
DNA is surrounded
by counter ions



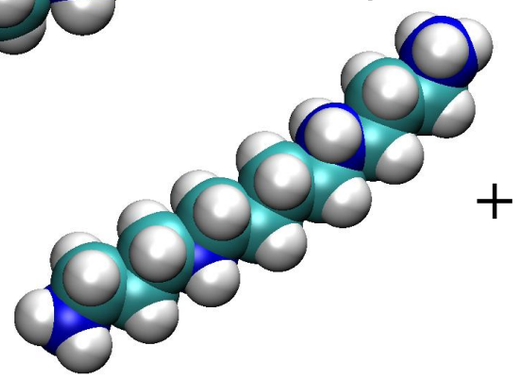
+1e, sodium or potassium



+2e, magnesium or calcium



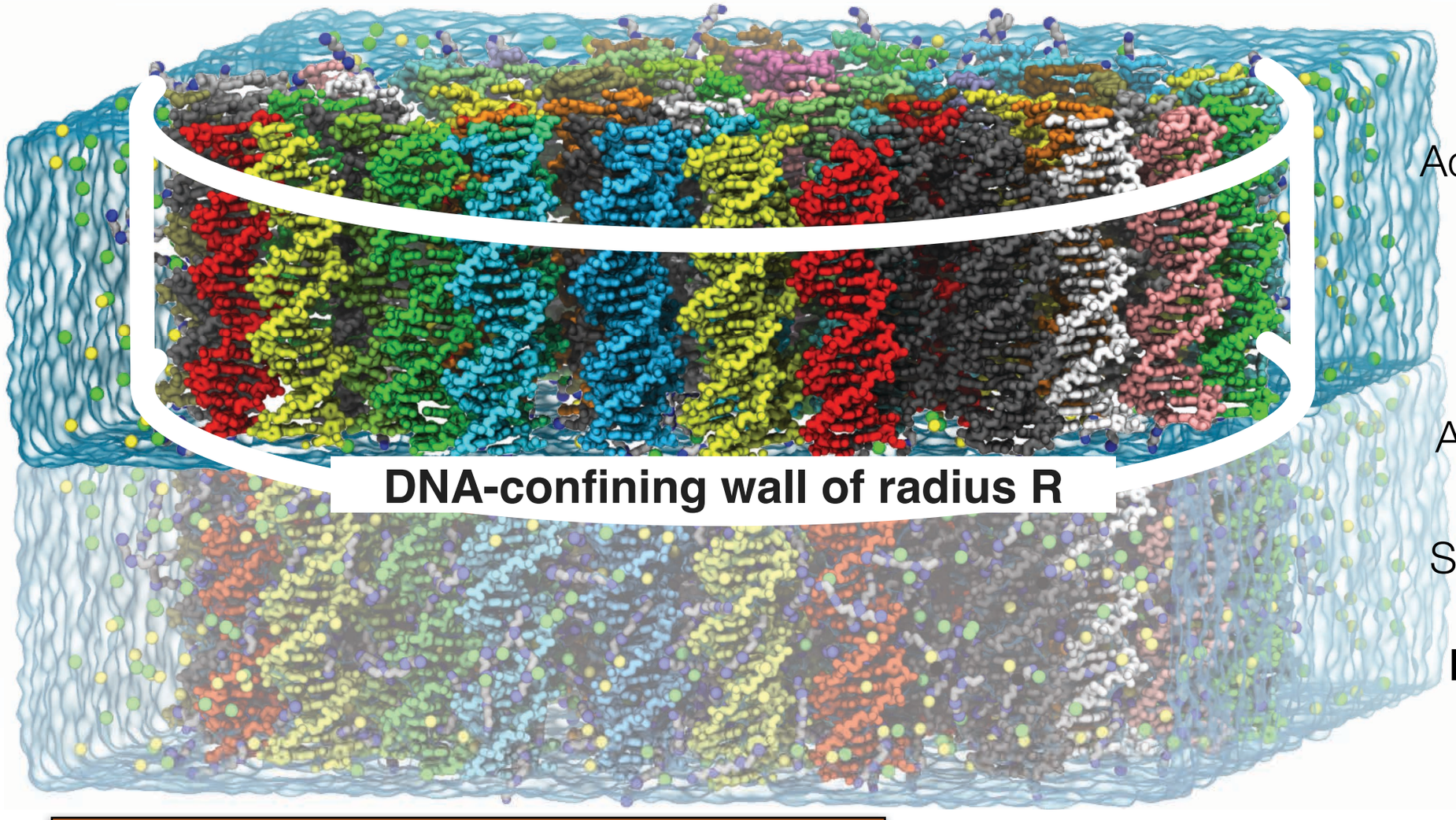
+3e,spermidine



+4e,spermine

Effective attraction between DNA is observed
when counterions have charge ≥ 2e

All-Atom Molecular Dynamics Simulation of DNA Condensates



Add **64 DNA** helices

Add **polyamine cations (+4)**

Add 150 mM **NaCl**

Add explicit **water**

Apply a half-**harmonic wall potential** only to DNA

Solve the equation of motion (F= ma) under **periodic boundary condition** in all directions

Classical **Force Field**

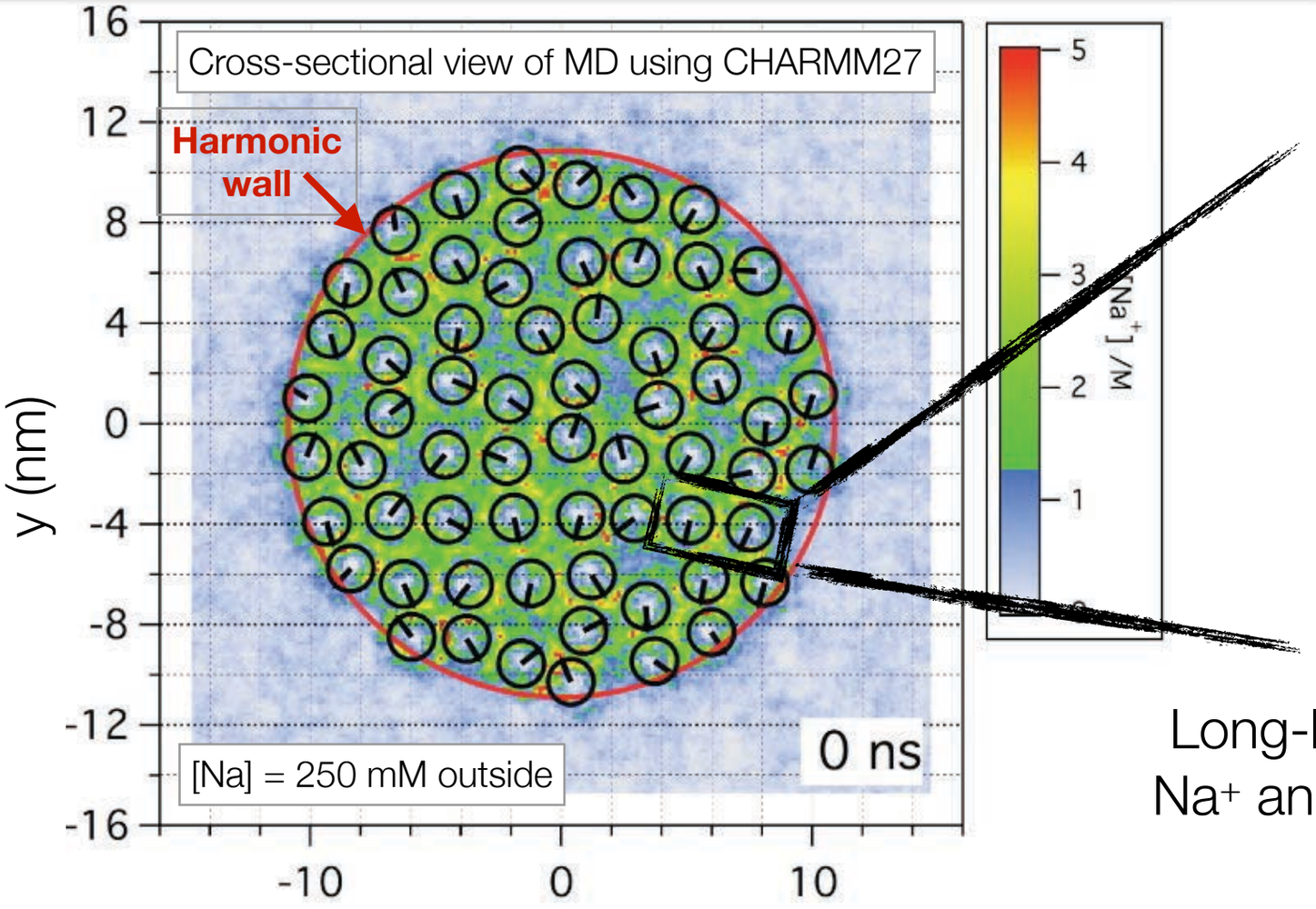
$$U(r) = \sum_{\text{bonds}} k_b(b - b_0)^2 + \sum_{\text{angles}} k_\theta(\theta - \theta_0)^2 + \sum_{\text{dihedrals}} k_\phi(1 + cos(n\phi - \phi_0)) + \sum_{\text{non-bonded pairs } i,j} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} + \sum_{\text{non-bonded pairs } i,j} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]$$

Partial charges from quantum mechanics

LJ parameters from experiments

Bonded parameters from quantum mechanics

Standard CHARMM & AMBER Force Fields Are Not Perfect for the Simulation of DNA Condensates

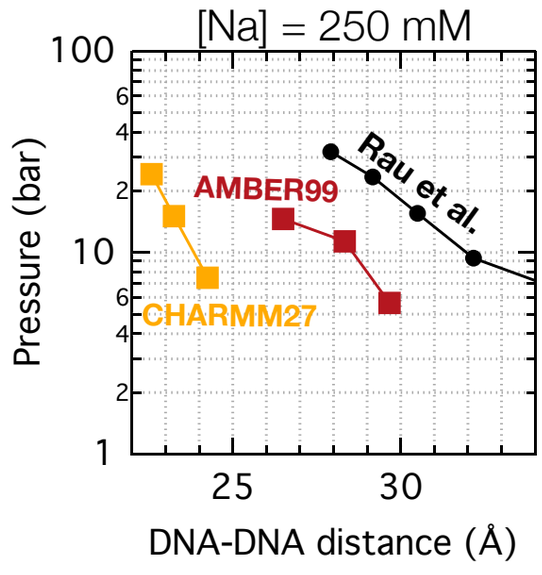


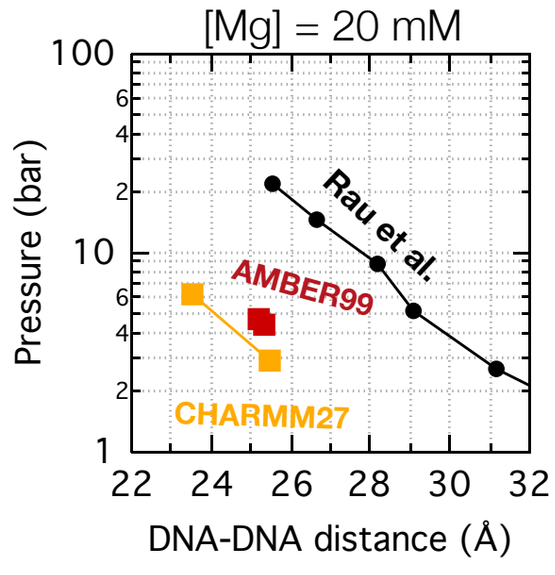
Na

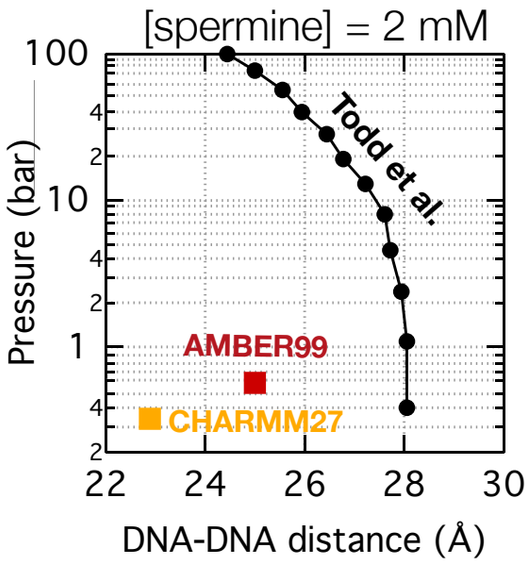
15-ns MD

Long-lasting **contact ion pairs** (CIP) between Na^+ and phosphate stabilize contact DNA pairs.

Due to excessive CIP formation, the simulations underestimate both inter-DNA distance and pressure in DNA array systems.



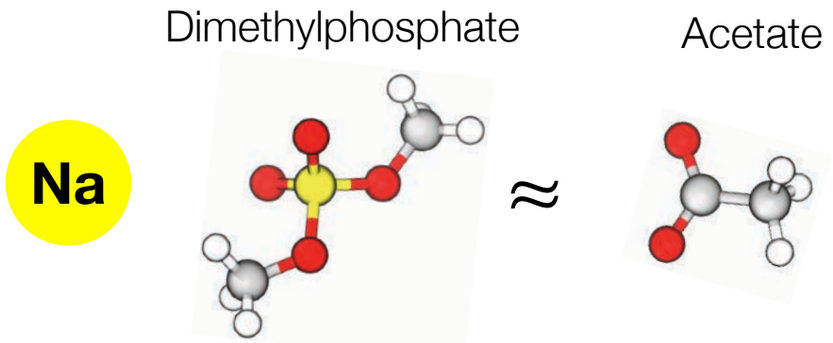


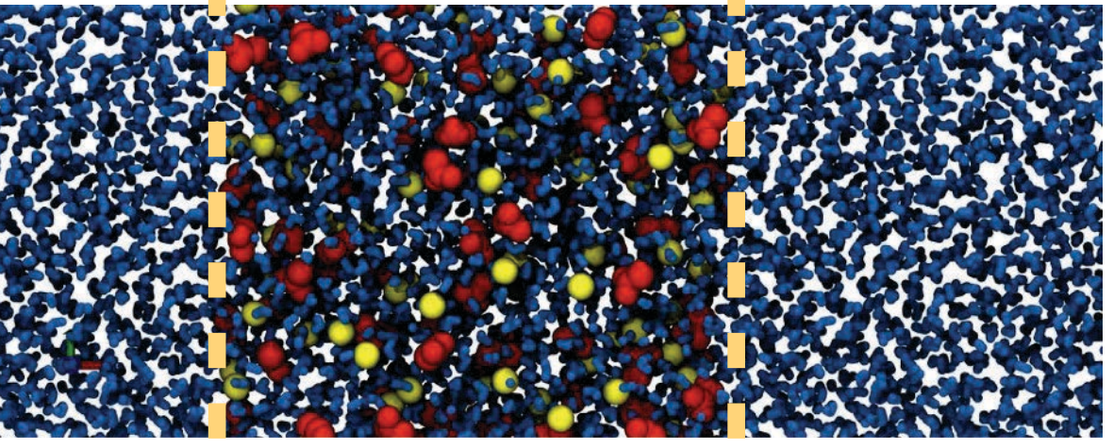


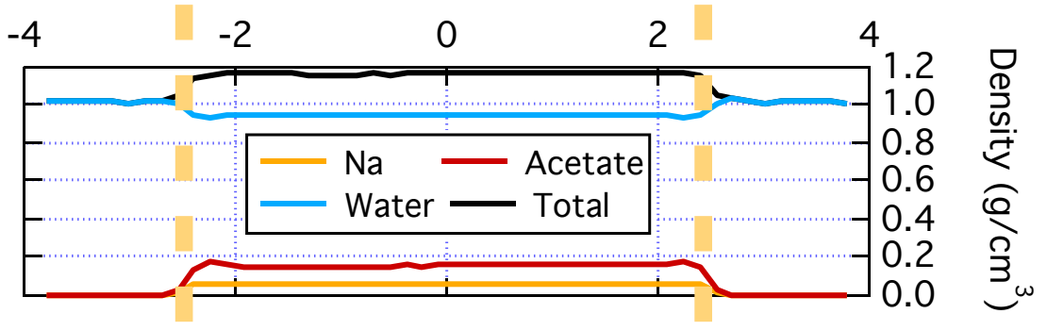
Rau et al, *PNAS* 1984 Todd et al, *BJ* 2008 Yoo & Aksimentiev, *JPCL* 2012

Champaign-Urbana Non-Bonded FIX (CUFIX): Improved Lennard-Jones Parameters for CHARMM & AMBER

"Much of what is known about association and dissociation of solutes and ions comes from measurements of **colligative properties**" — Molecular driving forces by Dill & Bromberg.

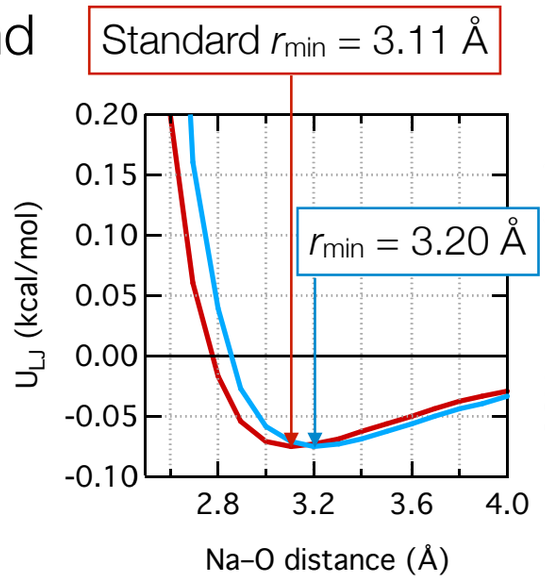


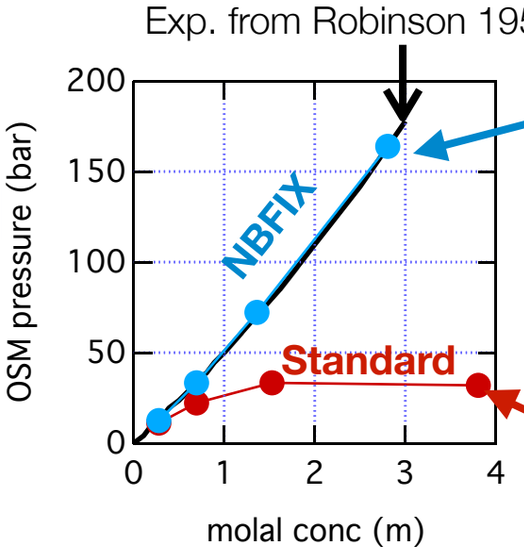


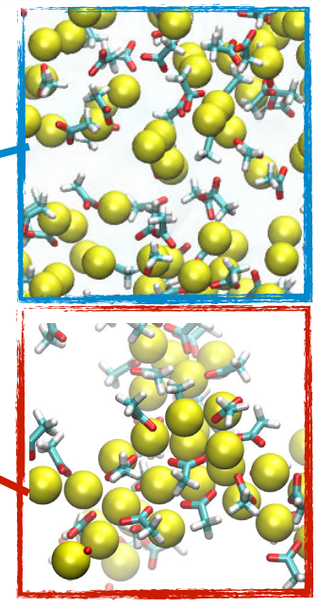


Yoo & Aksimentiev, *JPCL* 2012

Murad & Powles, *JCP* 1993
Luo & Roux, *JPCL* 2010

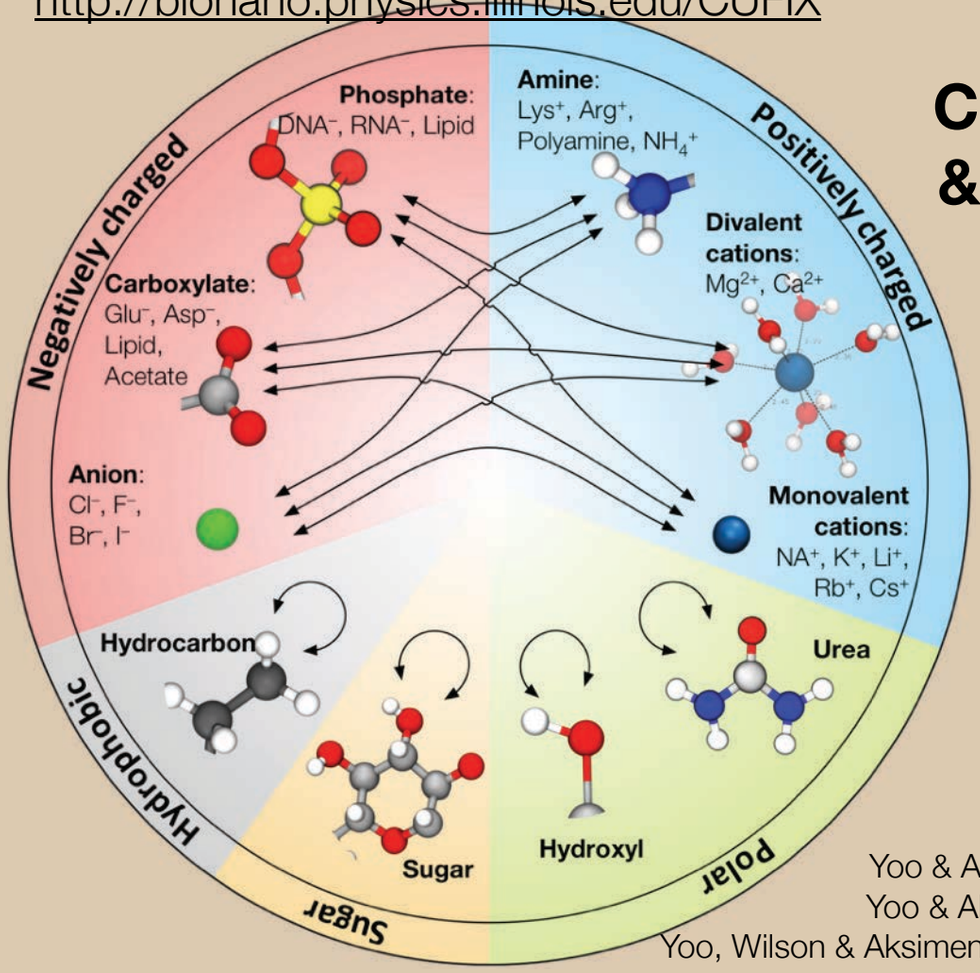






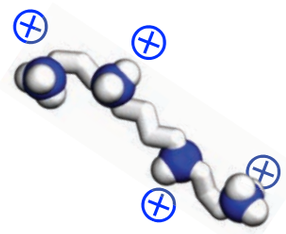
http://bionano.physics.illinois.edu/CUFIX

CUFIX for CHARMM36 & AMBER99

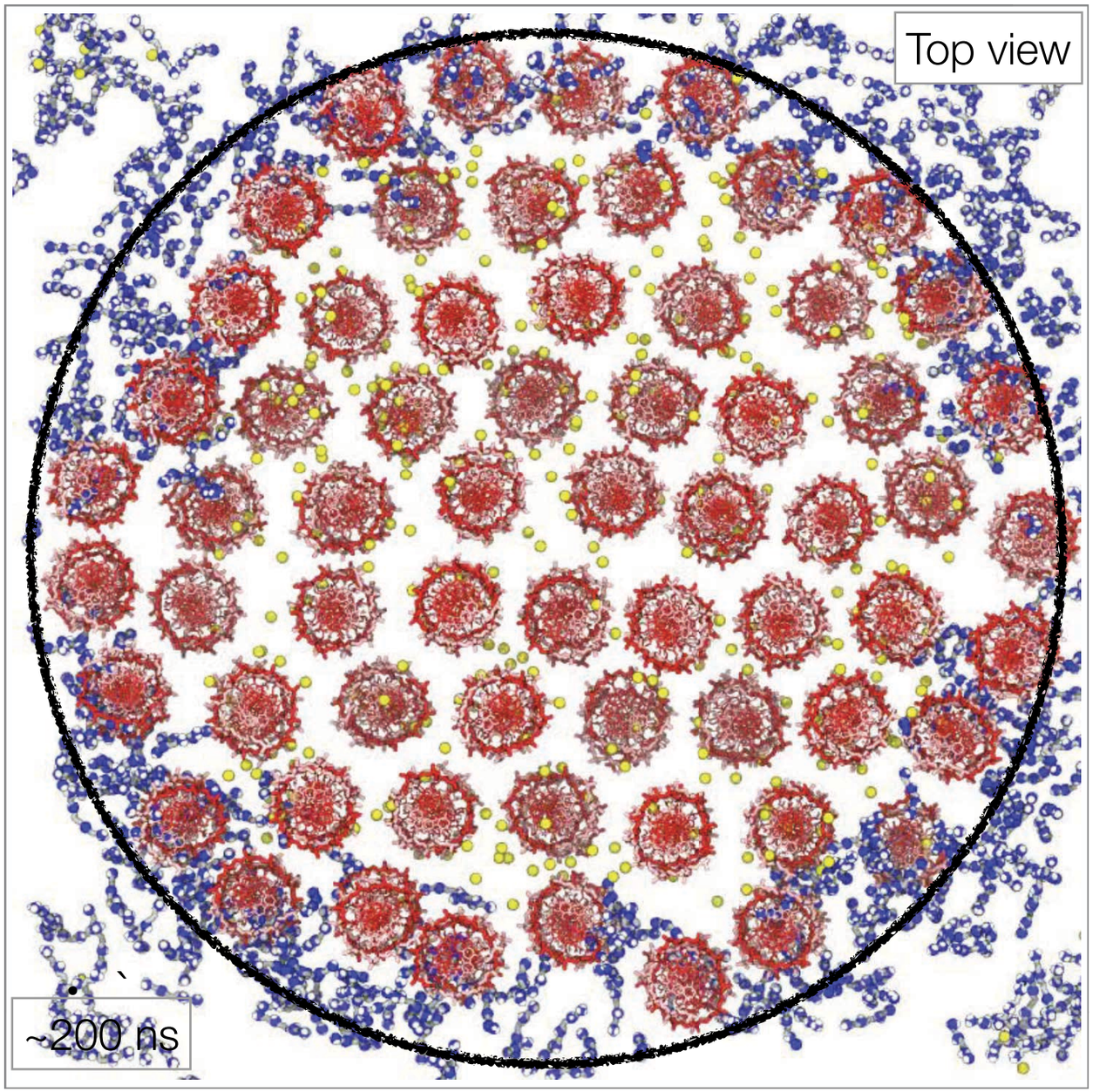


Yoo & Aksimentiev, *JPCL* 2012
Yoo & Aksimentiev, *JCTC* 2016
Yoo, Wilson & Aksimentiev, *Biopolymers* 2016

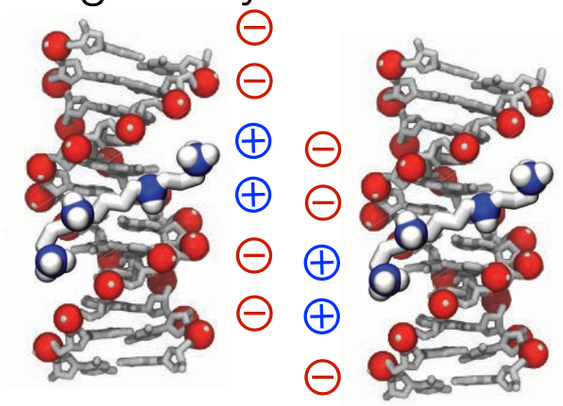
Bridging Ions Govern DNA Condensation

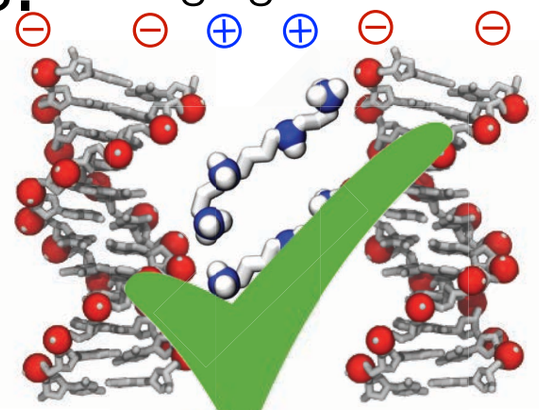


30 bar → 1 bar

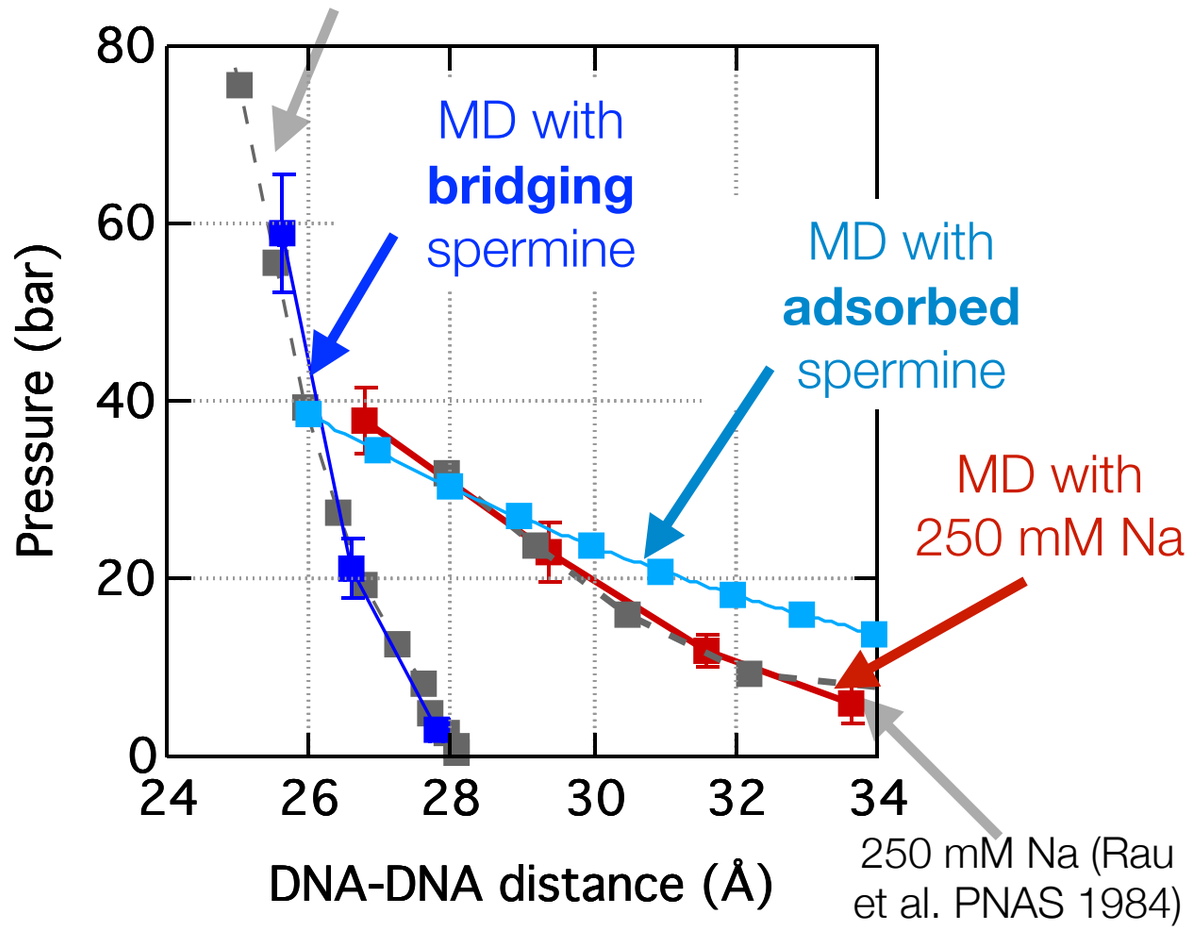


Wigner crystal model vs. Bridging model





2 mM spermine
(Todd et al. BJ 2008)

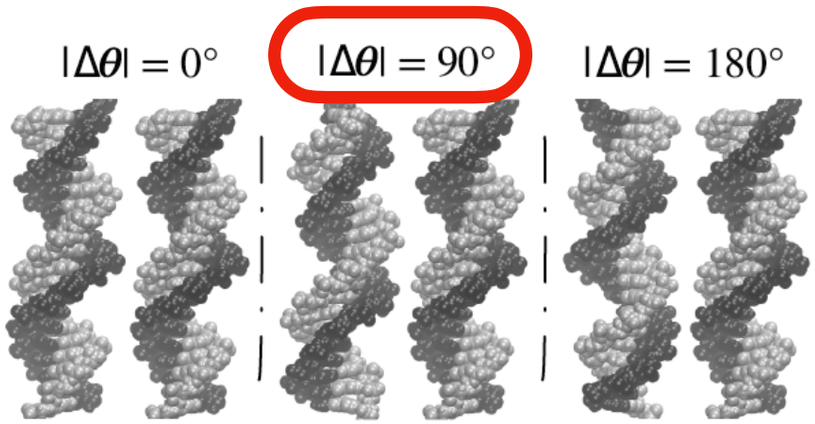


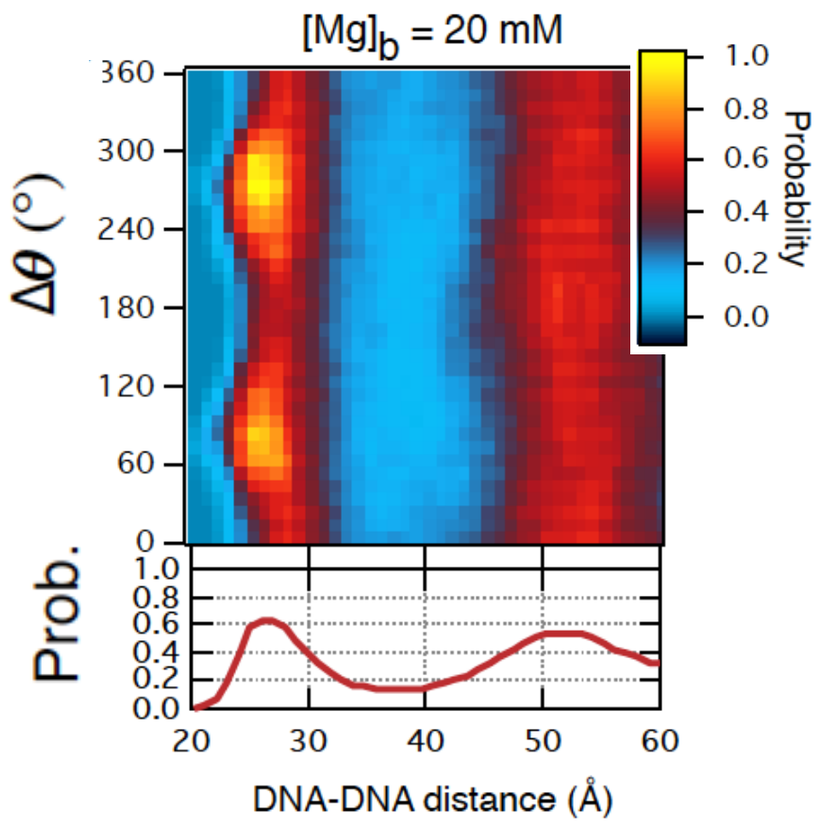
Yoo & Aksimentiev, *NAR* 2016

What else we learned from DNA array simulations

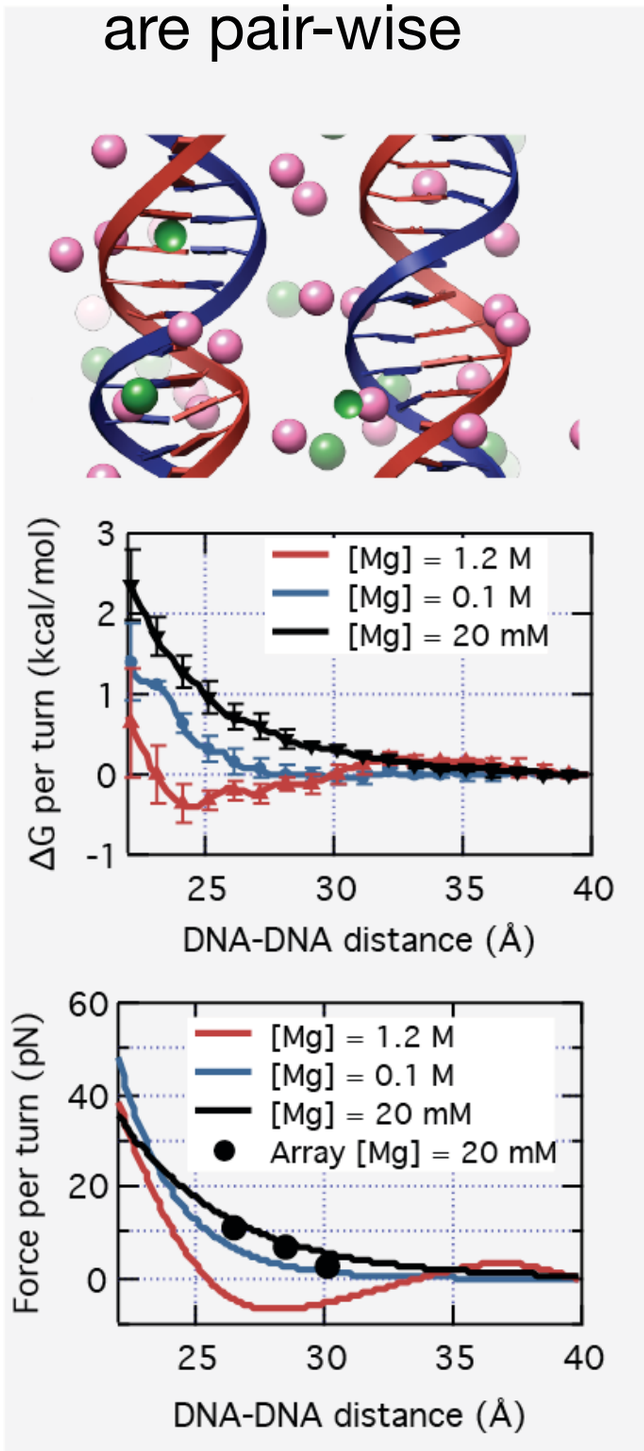
Yoo and AA, *Nucleic Acids Research* 44: 2036-2046 (2016)

Orientation of DNA helices is azimuthally correlated

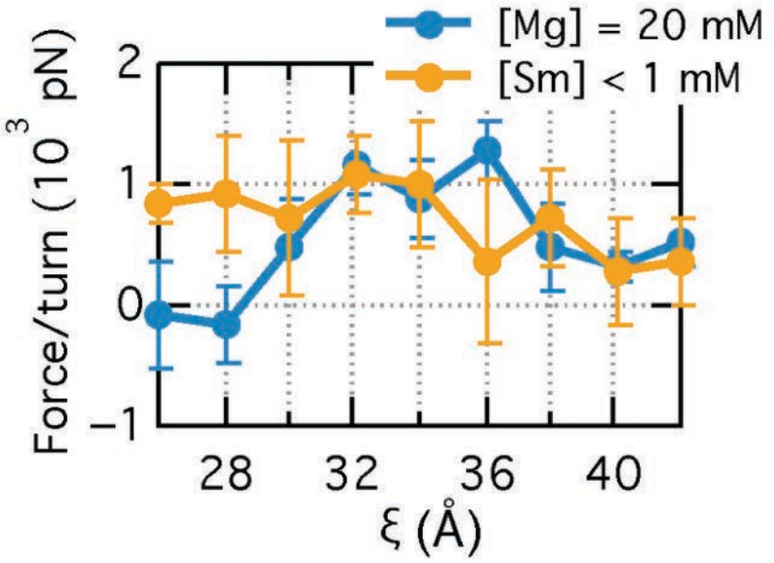




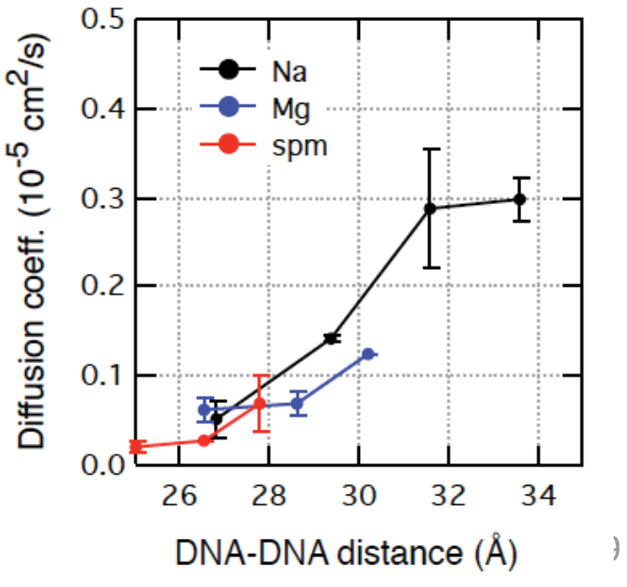
DNA-DNA forces in array are pair-wise



Electrostatics, not hydration produces DNA condensation

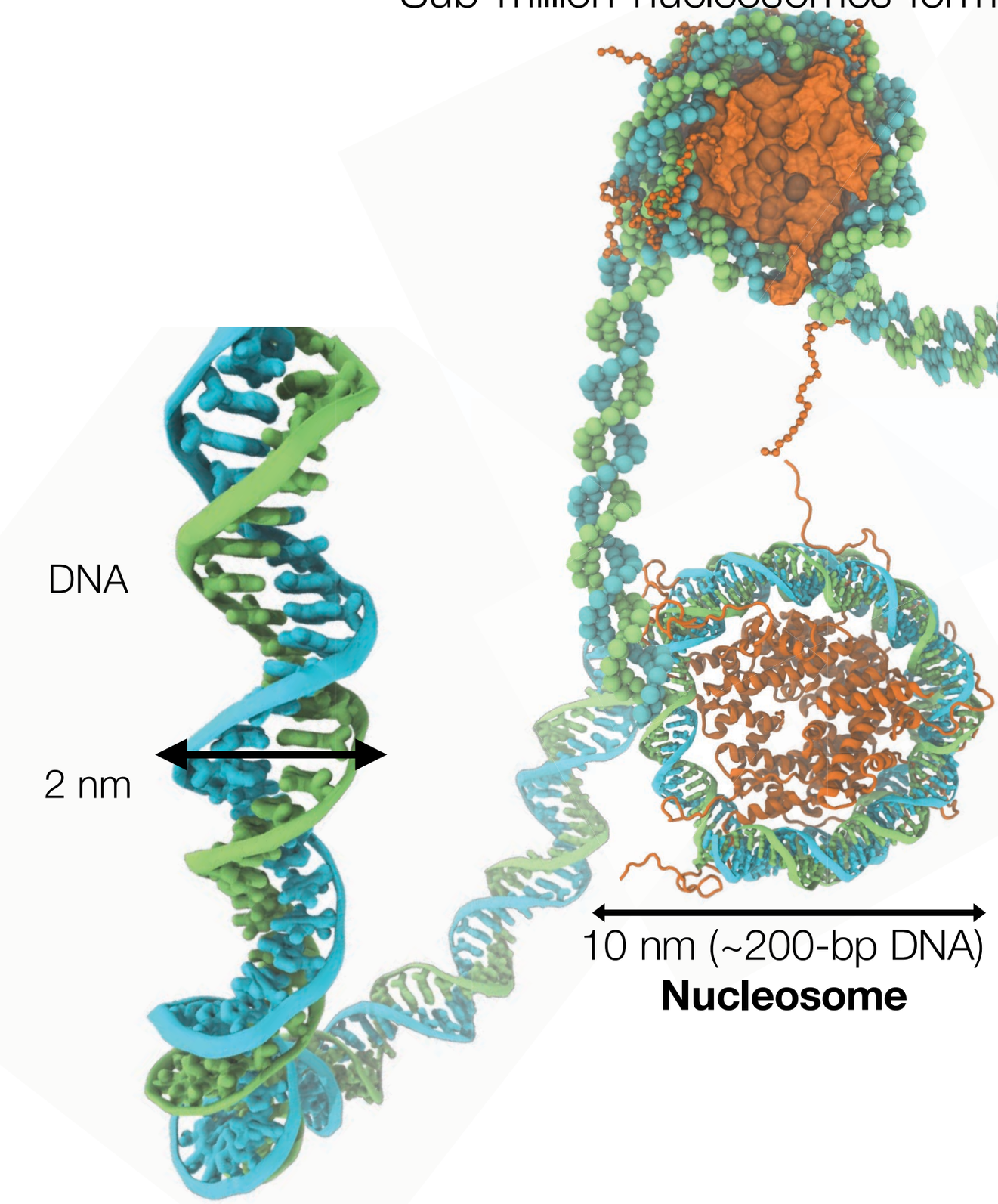


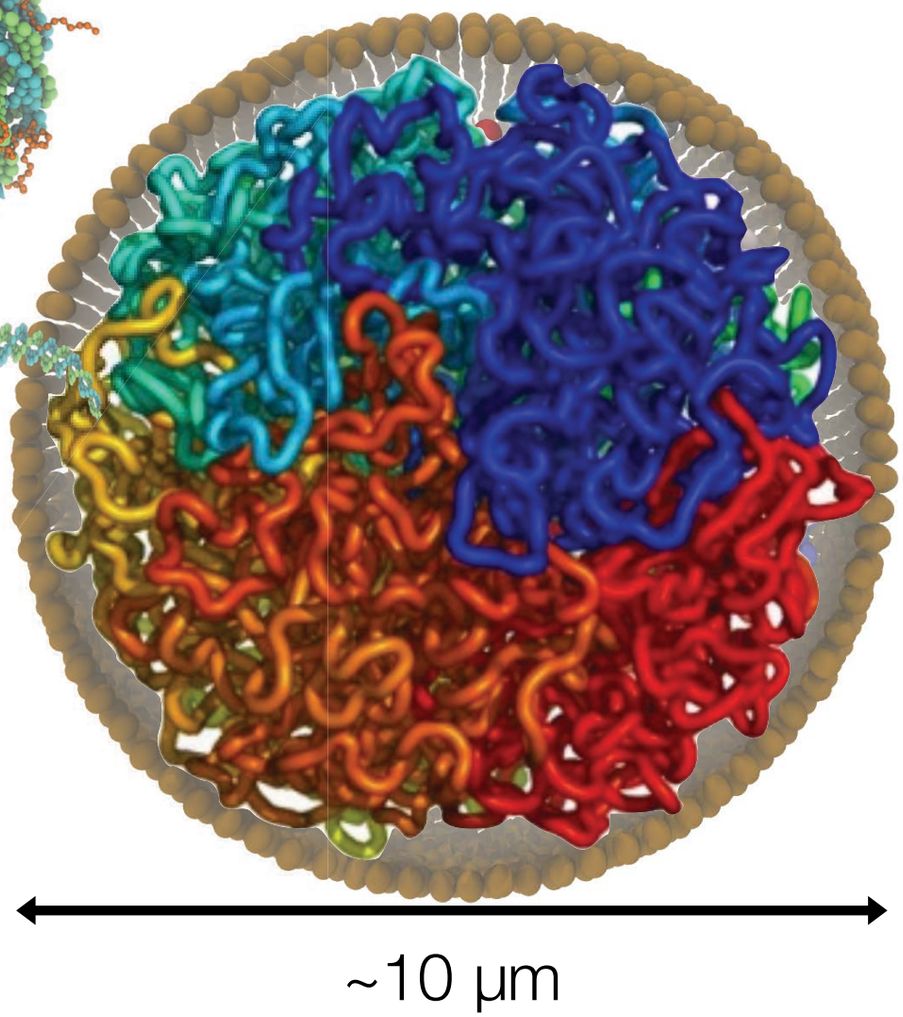
Inter-DNA friction depends mostly on DNA—DNA distance



3D Organization of Human Chromosomes

Sub-million nucleosomes form a **chromatin** fiber.

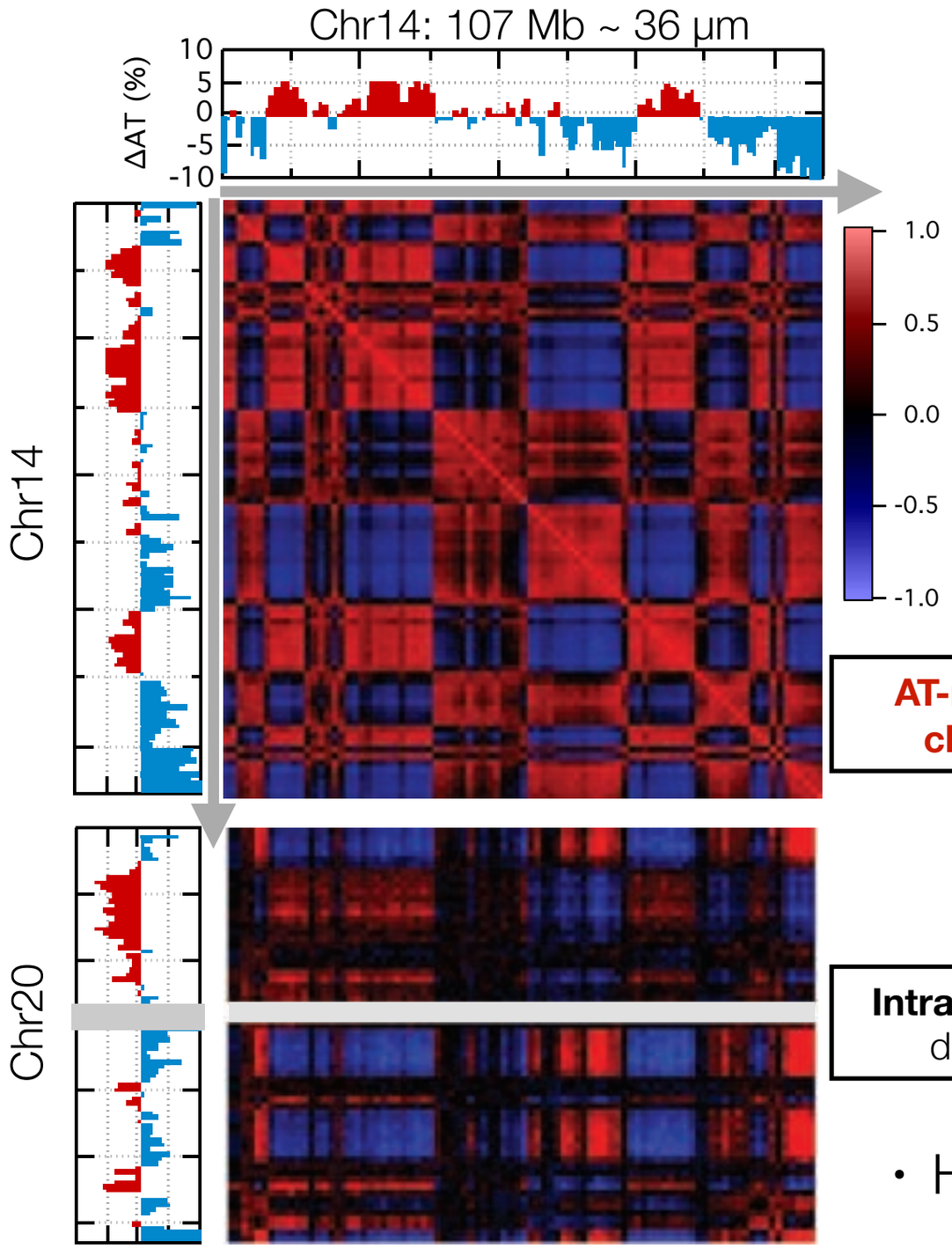




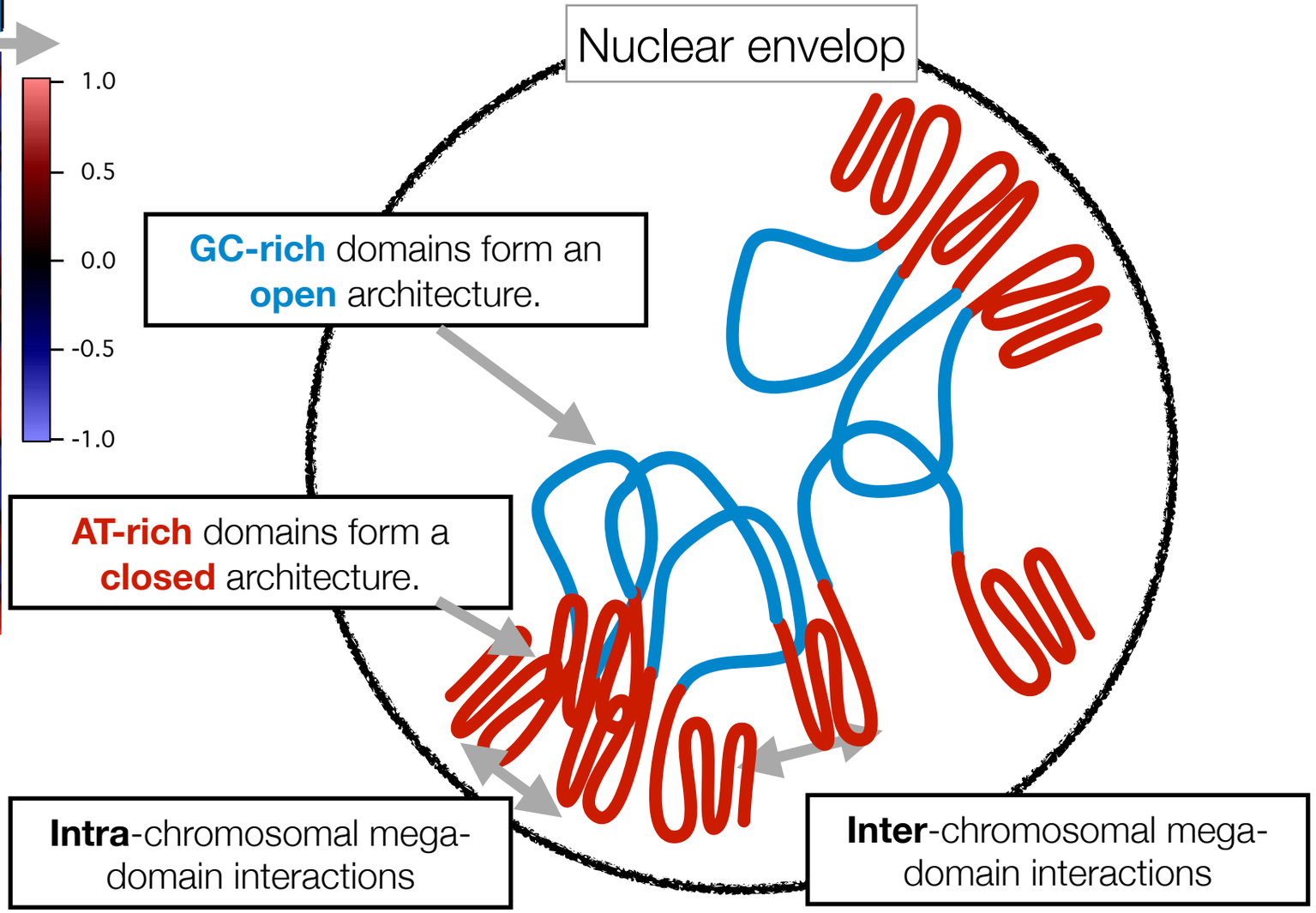
- How are the chromosomes organized?
- What are the fundamental forces that define the chromosome organizations?

Lieberman-Aiden et al., Science 2009

Chromosome's Mega-Domains Can Recognize One Another



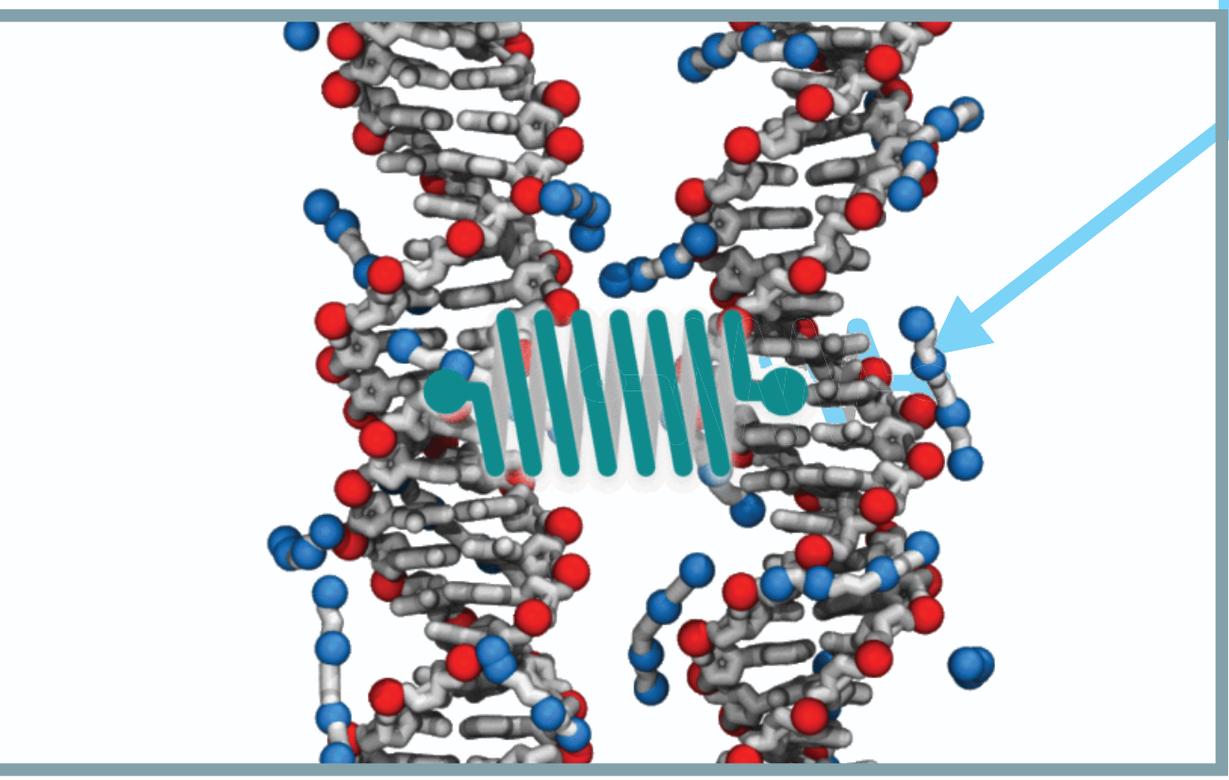
Hi-C contact correlation map revealed long-range contacts between AT-rich mega domains.



- How is such a long-range recognition possible?
- Can DNA recognize neighbors' AT contents without mediating proteins?

Lieberman-Aiden et al., *Science* 2009
Imakaev et al., *Nature Methods* 2012

AT Content Programs Strength of DNA Condensation

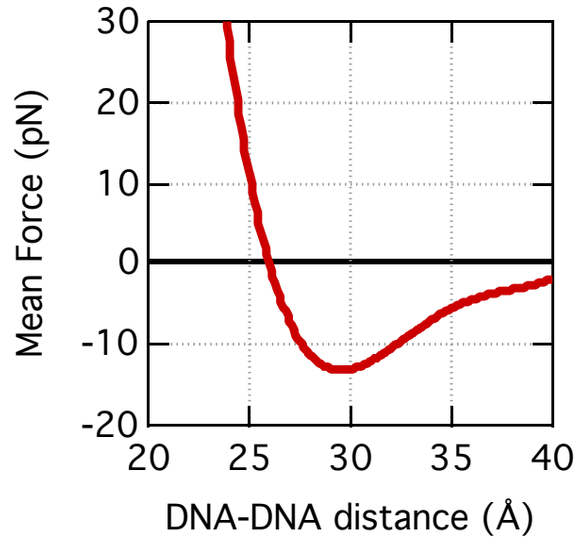


Harmonic spring connecting CM's reports the inter-molecular force.

$F = -kx$

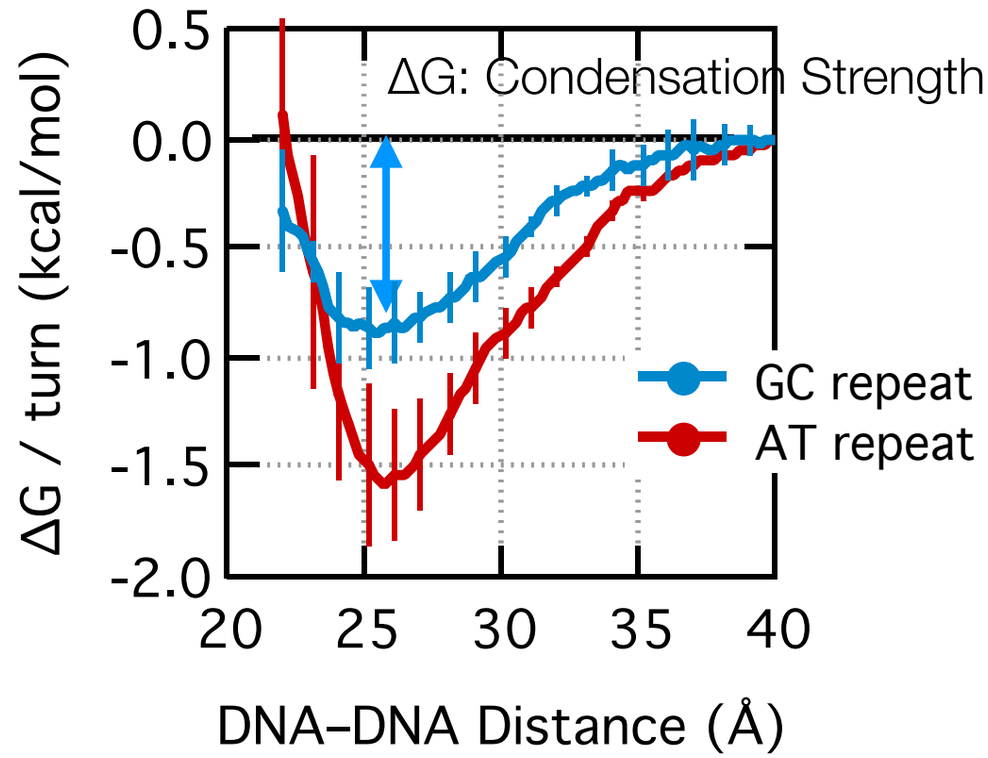
$\langle F(r) \rangle$

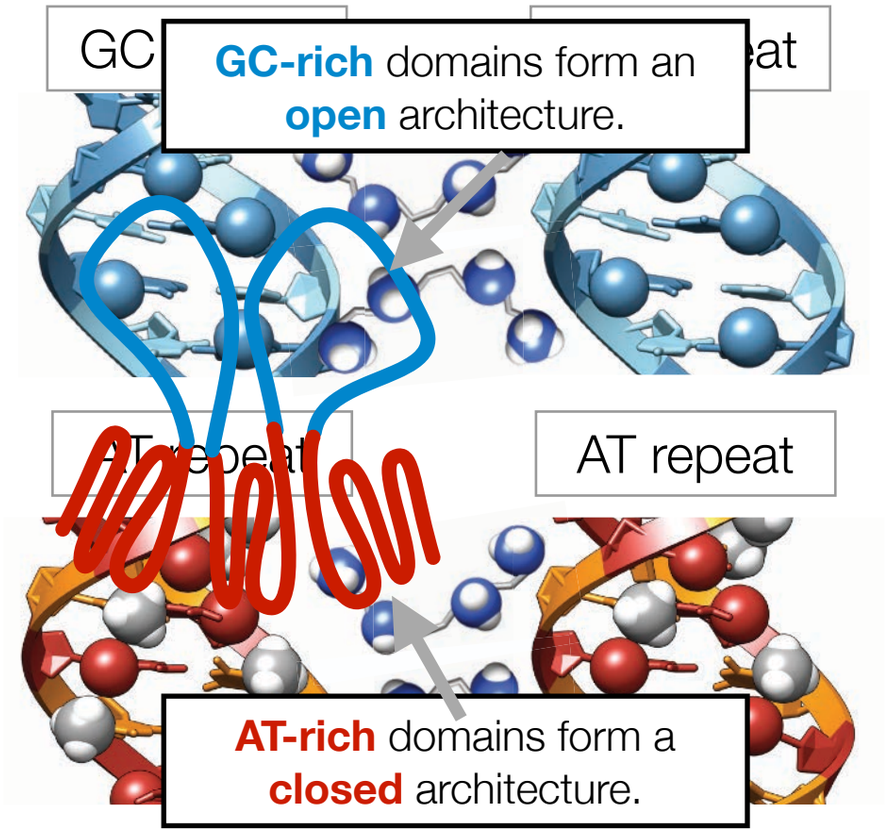
Compute Mean Force



Integrate Mean Force

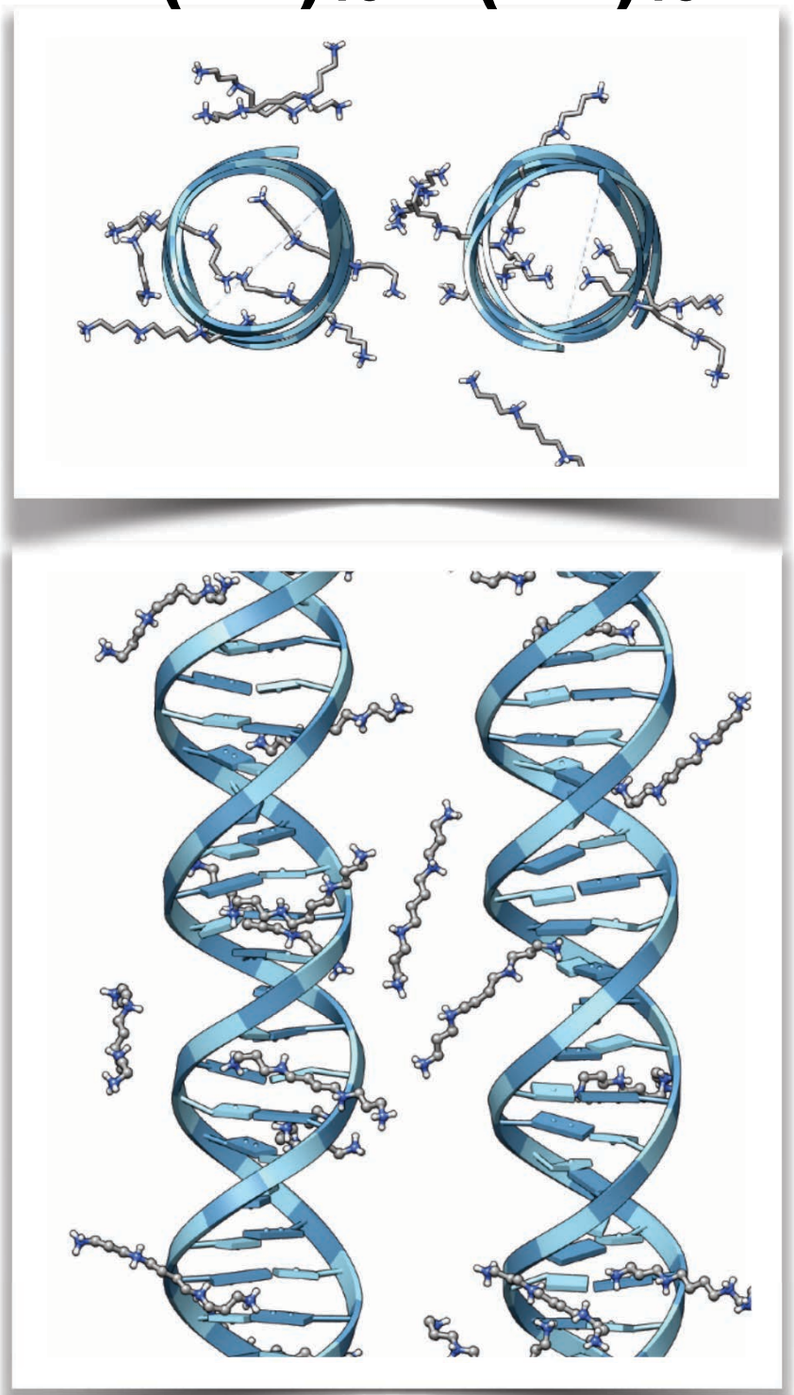
$$\langle F(r) \rangle = -\frac{d\Delta G}{dr}$$

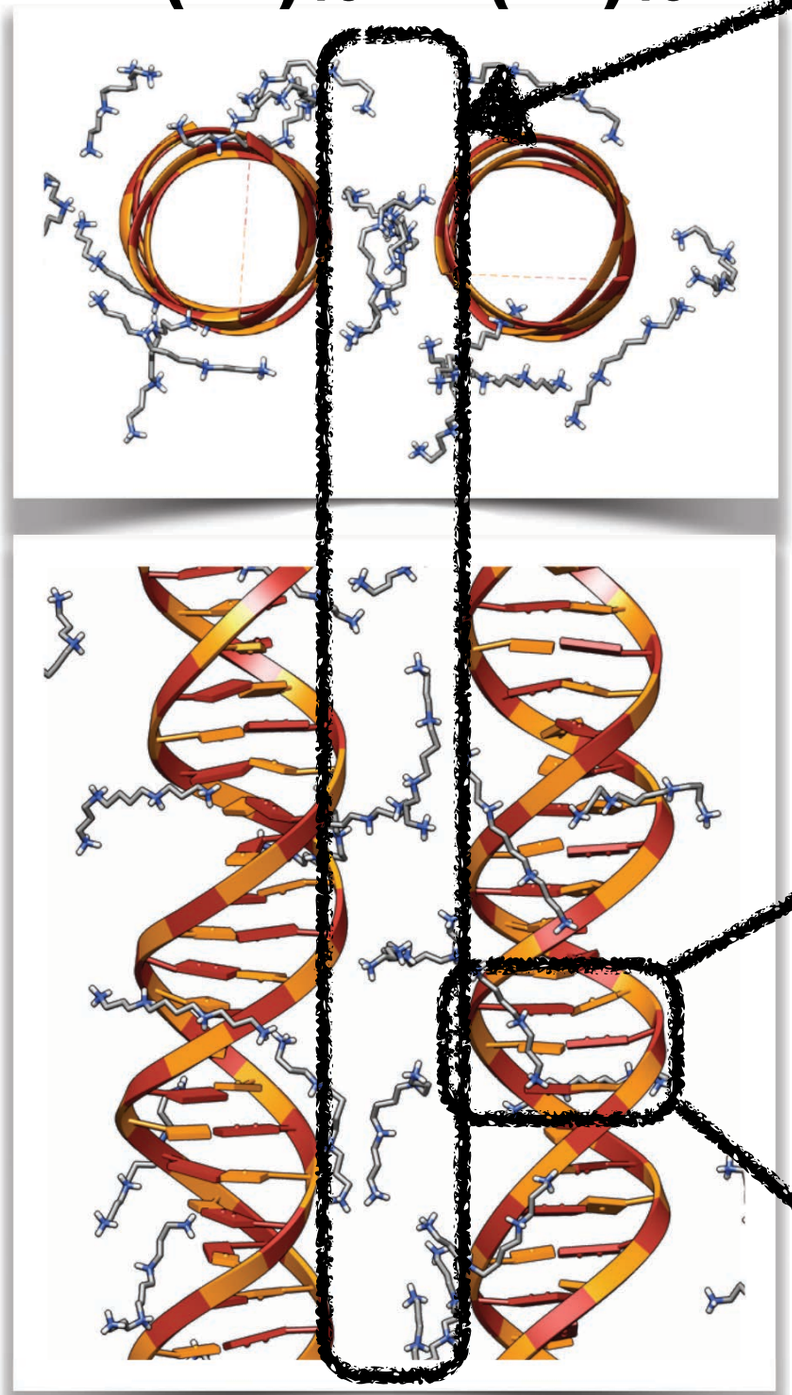




Yoo*, Kim*, Aksimentiev & Ha, *Nature Communications* (2016)

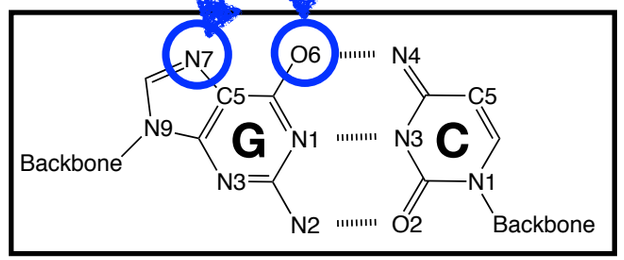
AT-rich segments form **clusters** better because they **share** polyamines with neighbors

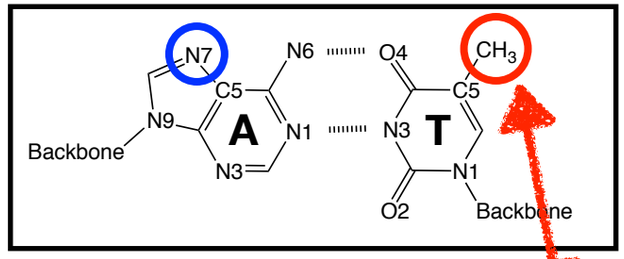




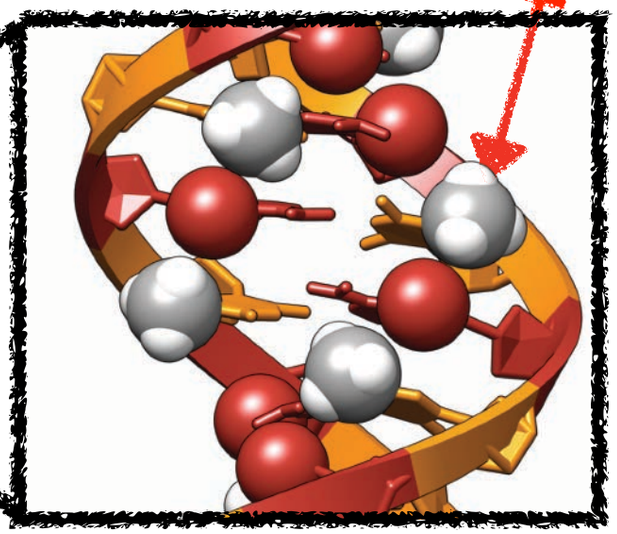
"Bridging" polyamines

Spermine binding site

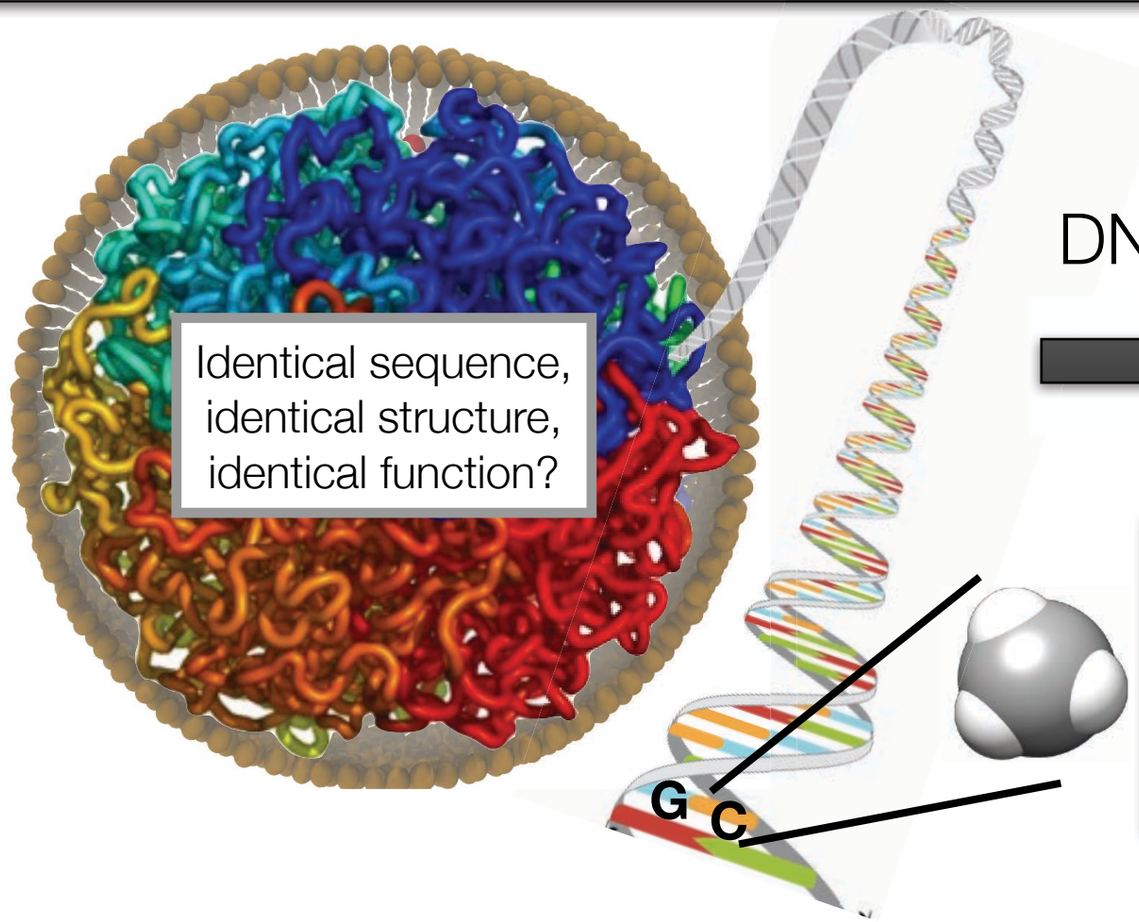




Methyl = Binding blocker

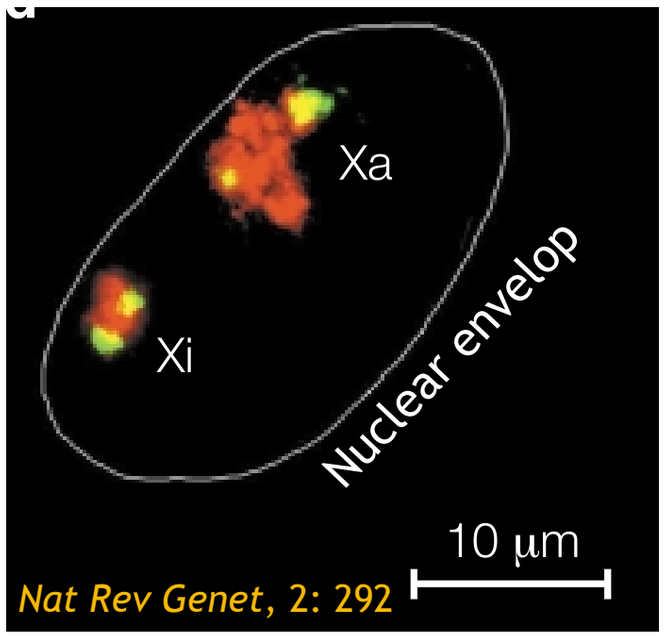


DNA Methylation Enhances Condensation



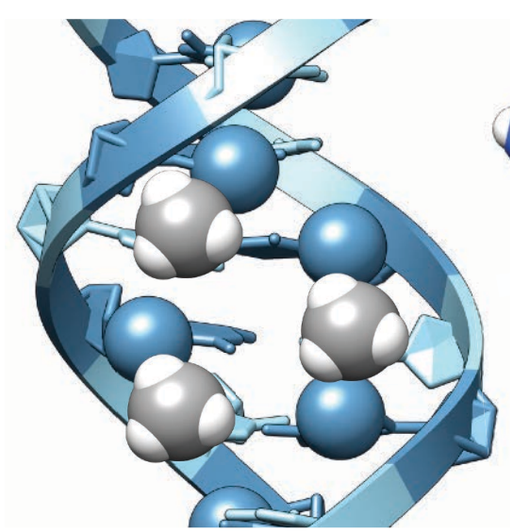
DNA **reprogramming**

Chemical addition of **methyl** group (**–CH_3**) to **cytosine** changes the structure of chromosome and the cellular functions.

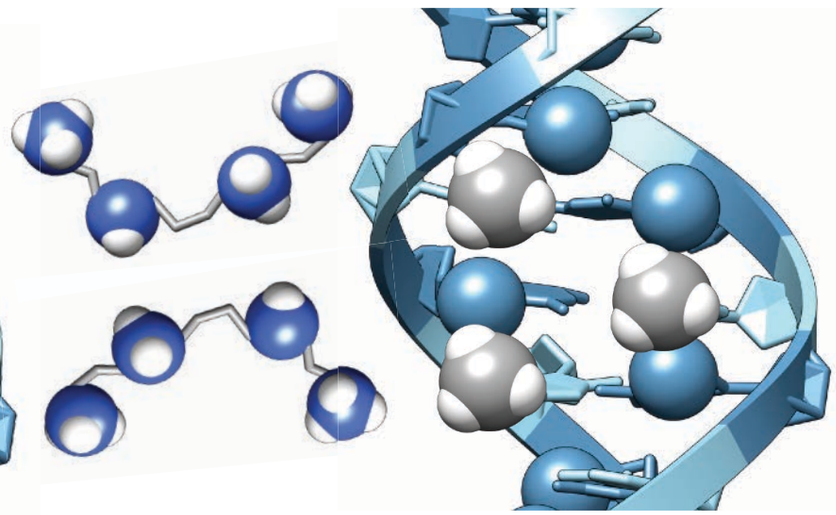


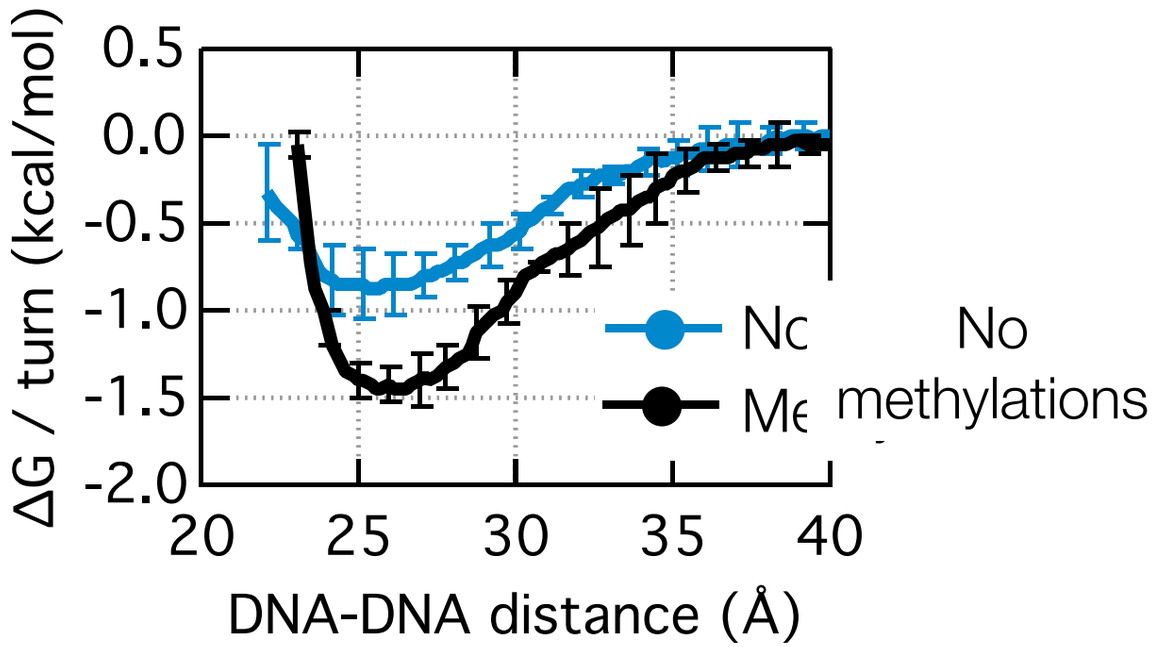
Highly methylated inactive X (Xi) chromosome is more compact than active X (Xa).

$[d^mCdG]_n$



$[d^mCdG]_n$



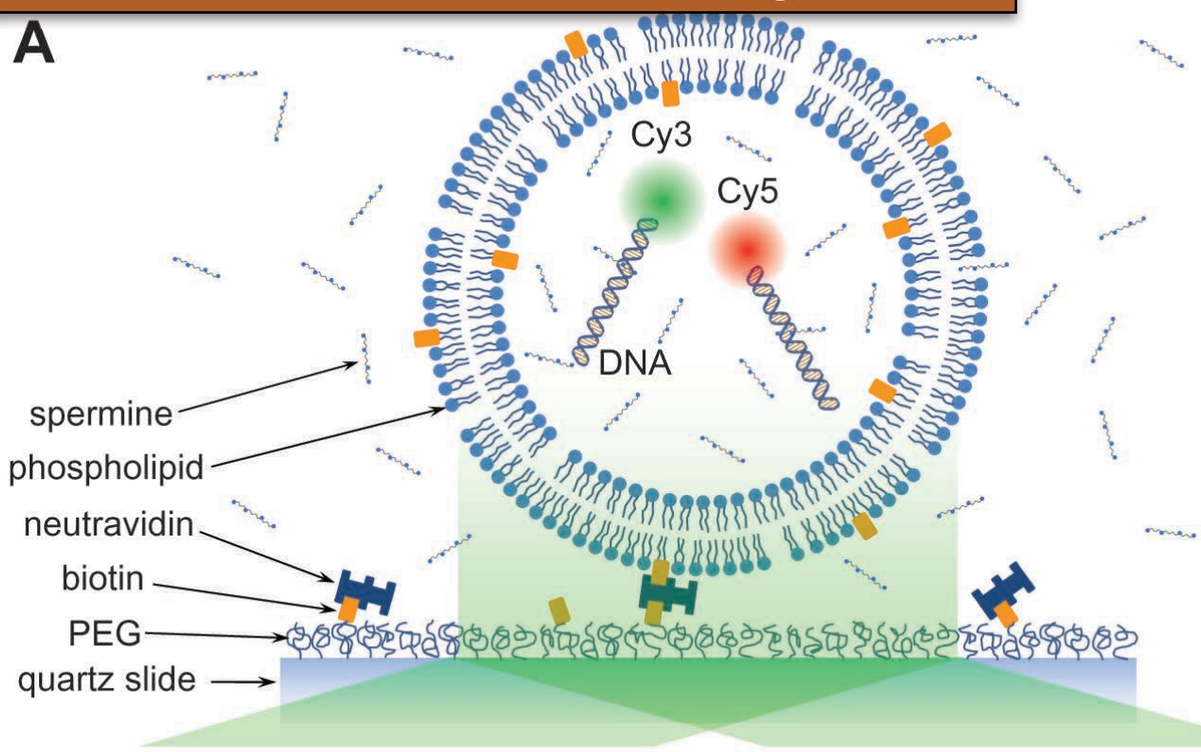


Yoo*, Kim*, Aksimentiev & Ha, *Nature Communications* (2016)

smFRET experiment confirms the prediction

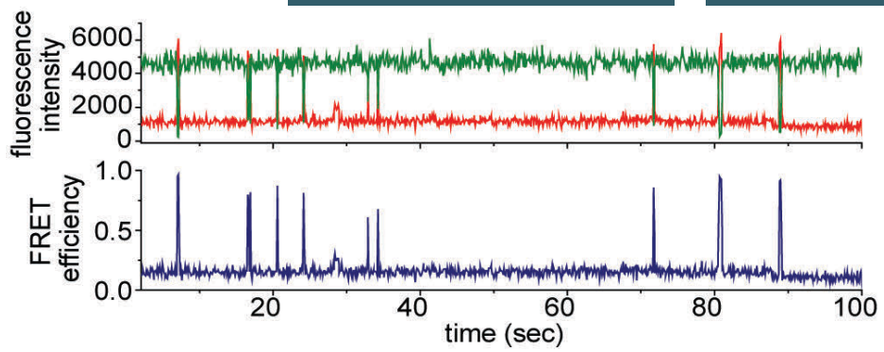
J. Yoo, H. Kim, A. Aksimentiev and T. Ha. *Nature Communications* 7:11045 (2016)

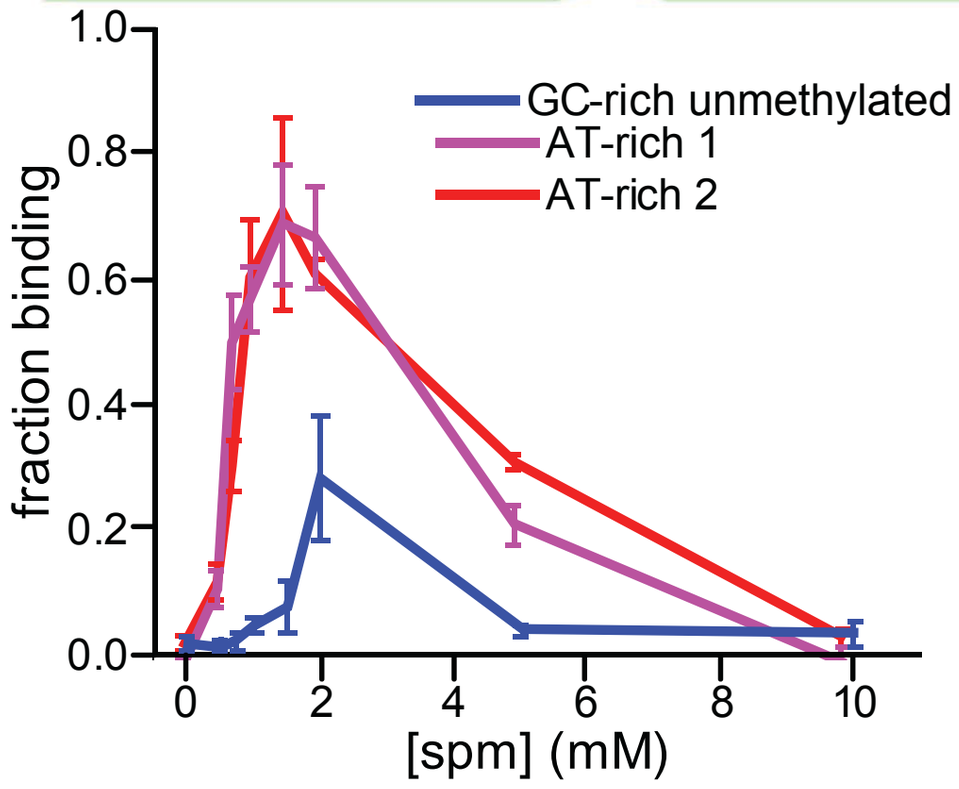
FRET: Fluorescence Resonance Energy Transfer

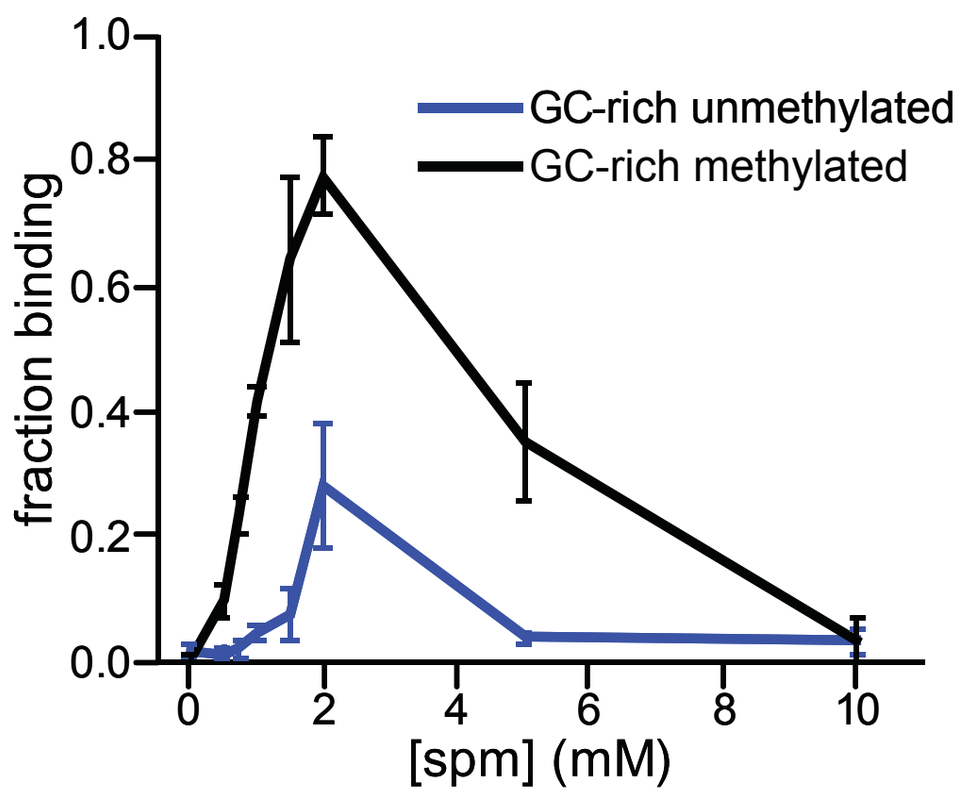


Hajin Kim

Taekjip Ha

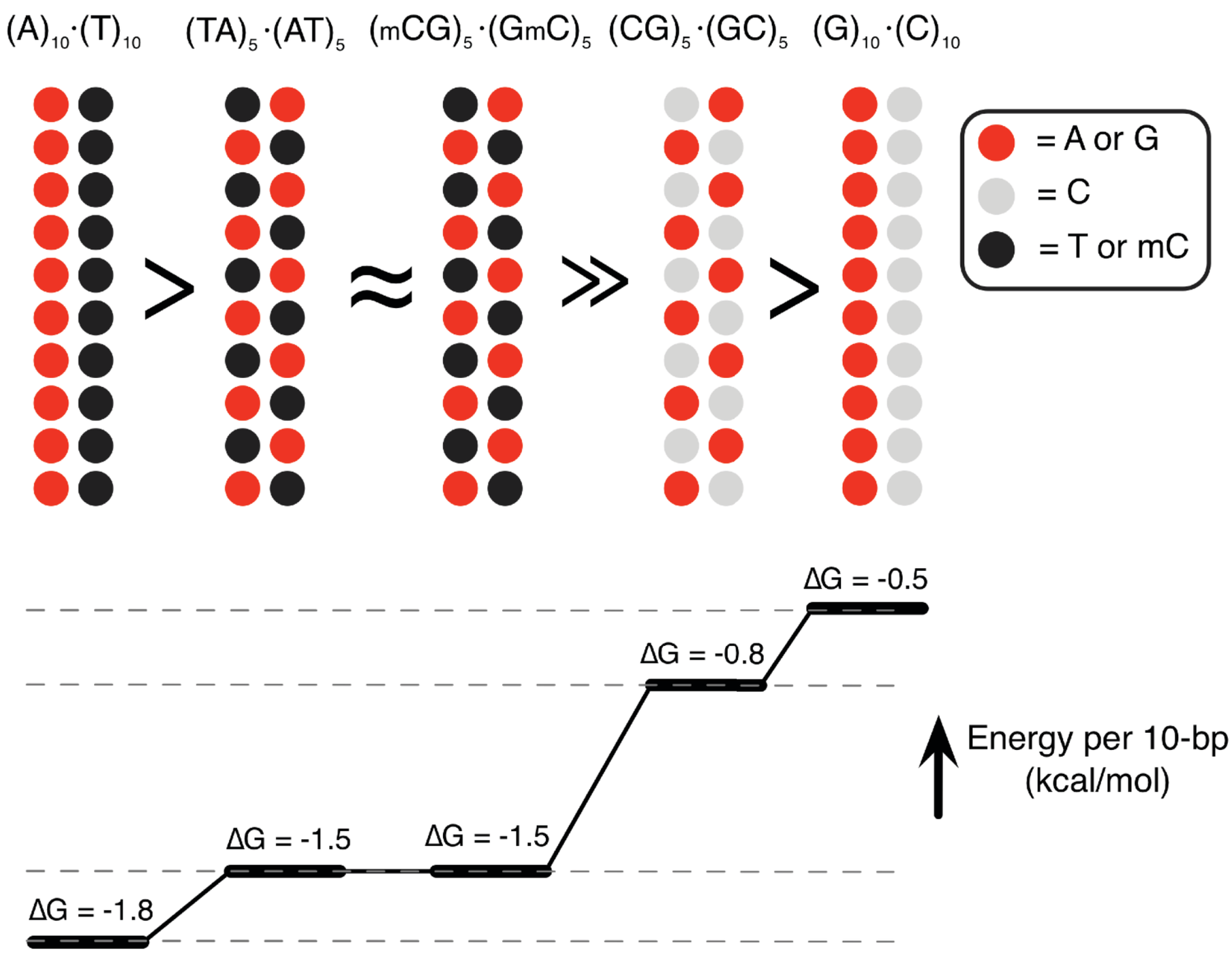




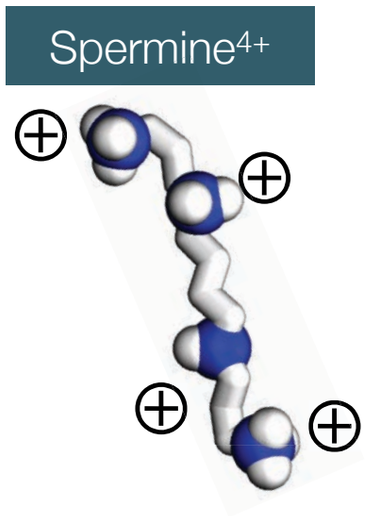


DNA attraction controlled by methylation pattern

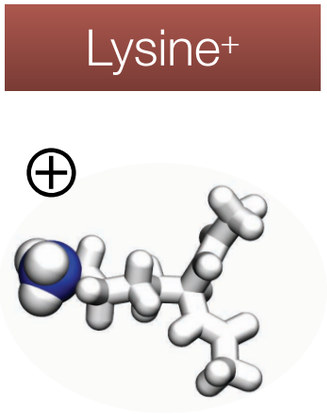
J. Yoo, H. Kim, A. Aksimentiev and T. Ha. Nature Communications 7:11045 (2016)

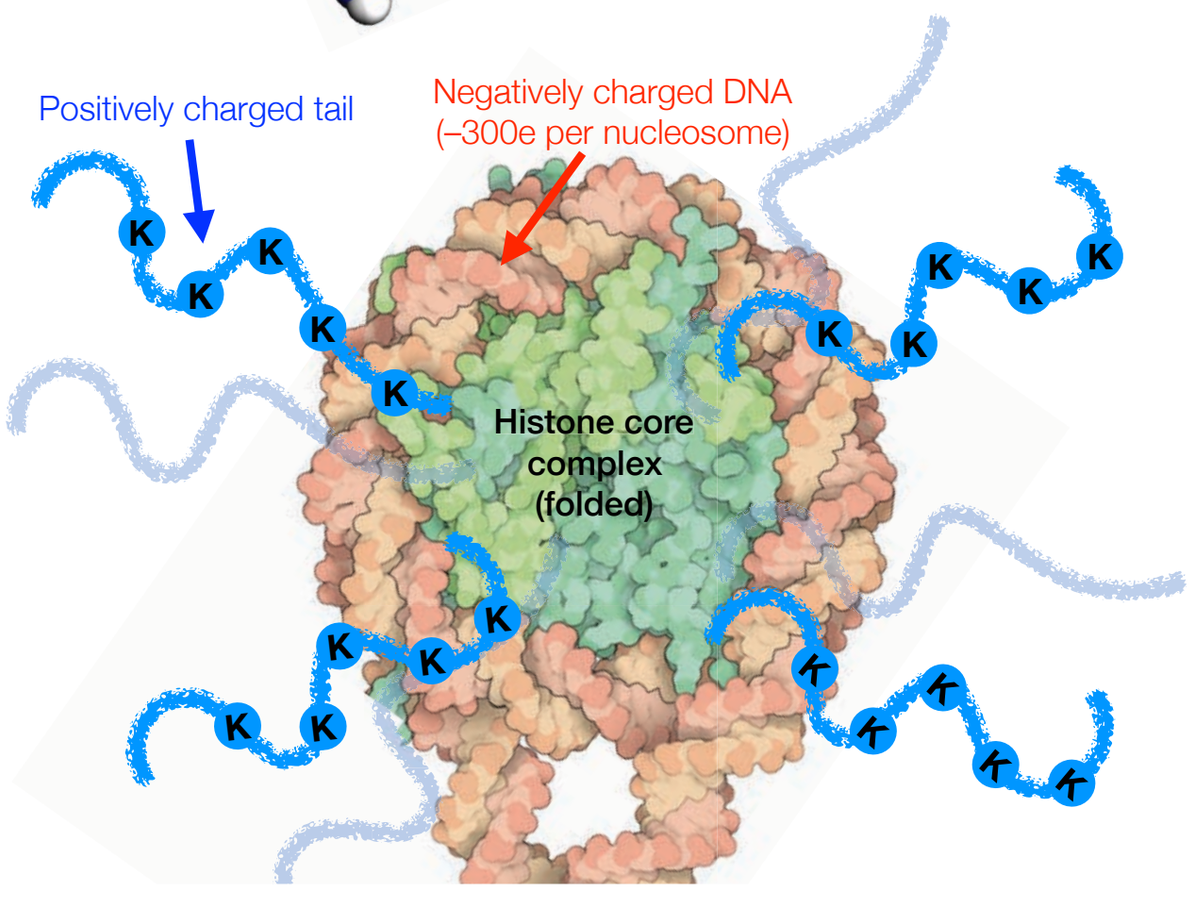


DNA condensation by **lysine polypeptides**

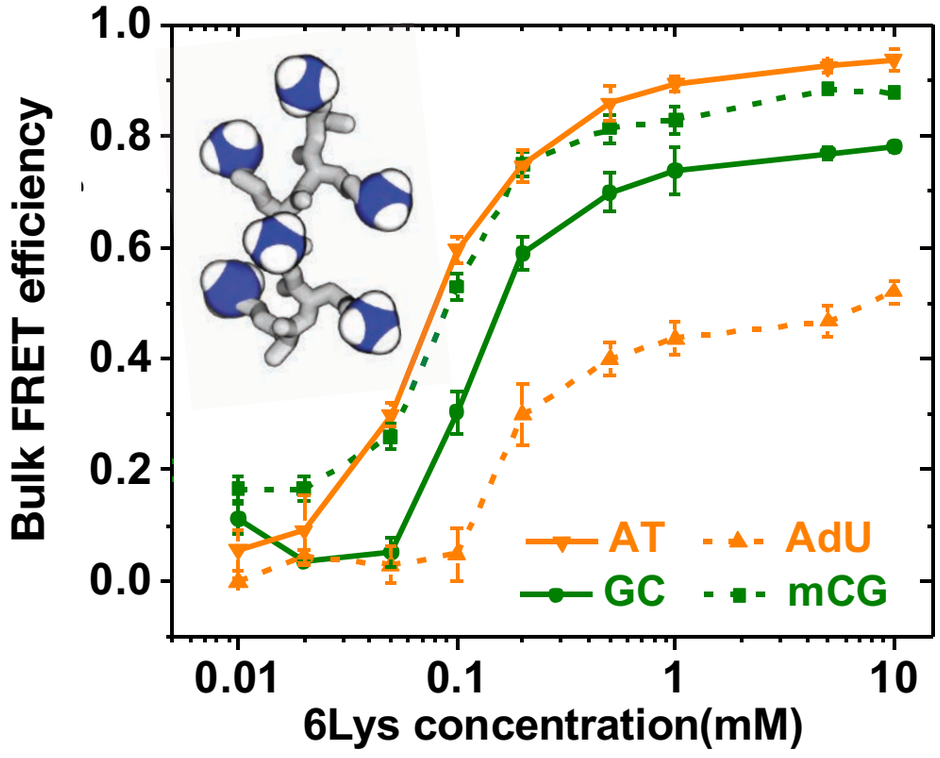


≈ ?



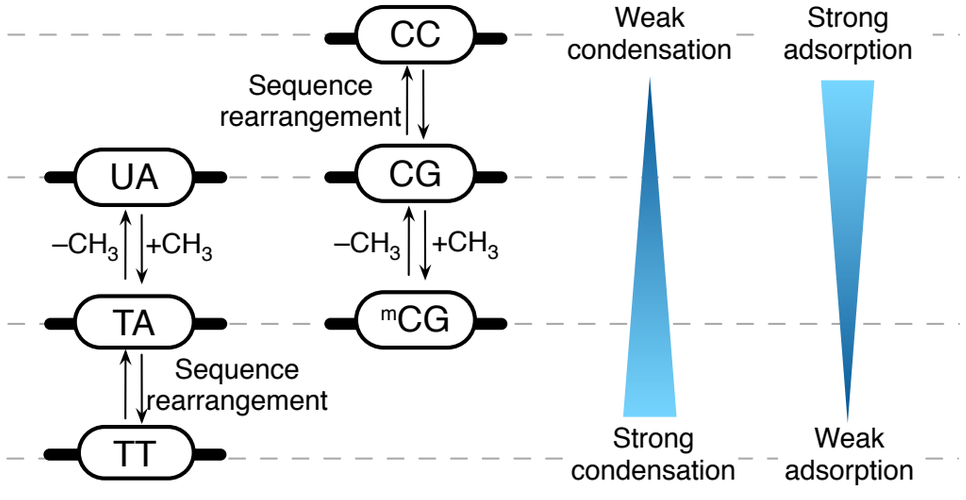


***Nucleic Acids Res.* 46: 9401 (2018)**



Hyunju Kang @ UNIST

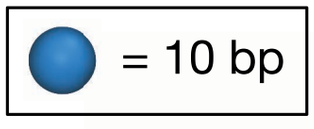
RNA has no methyl groups and is more difficult condense. Tolokh et al., *Nucleic Acids Research* (2014)

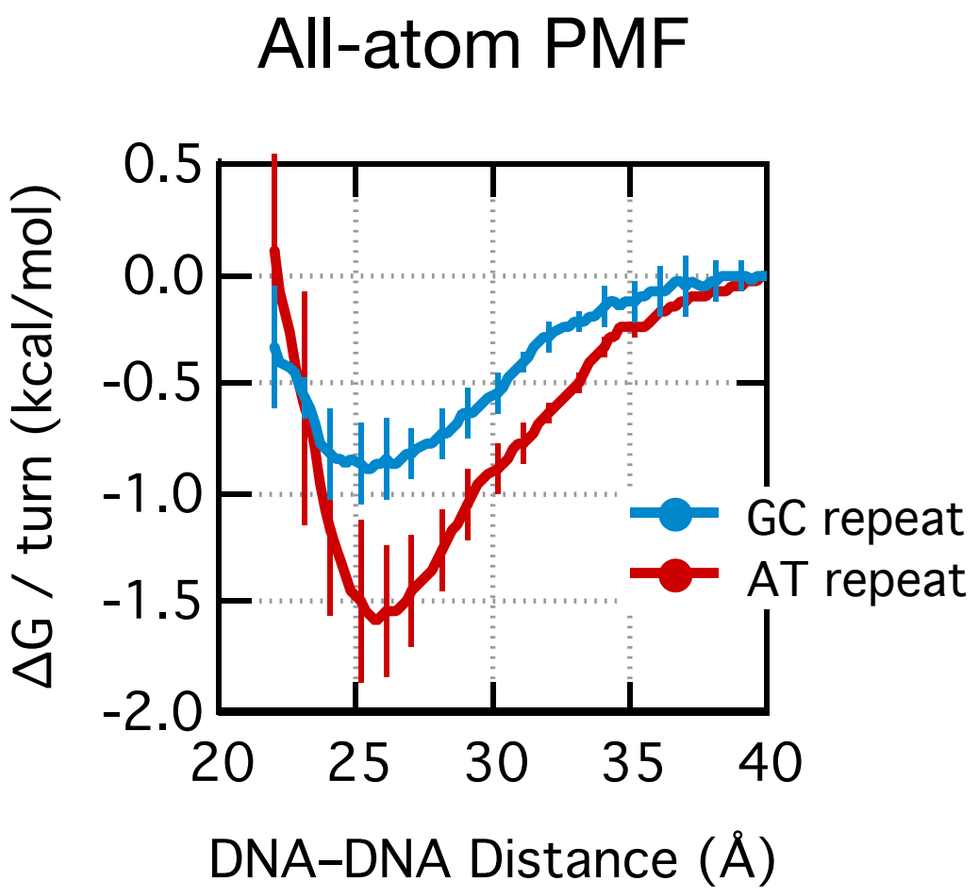


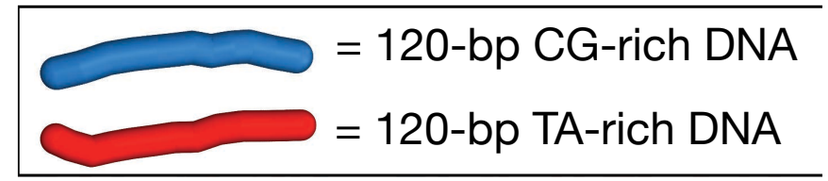
Using **lysine peptides** instead of spermine does not change the sequence-dependence of DNA condensation

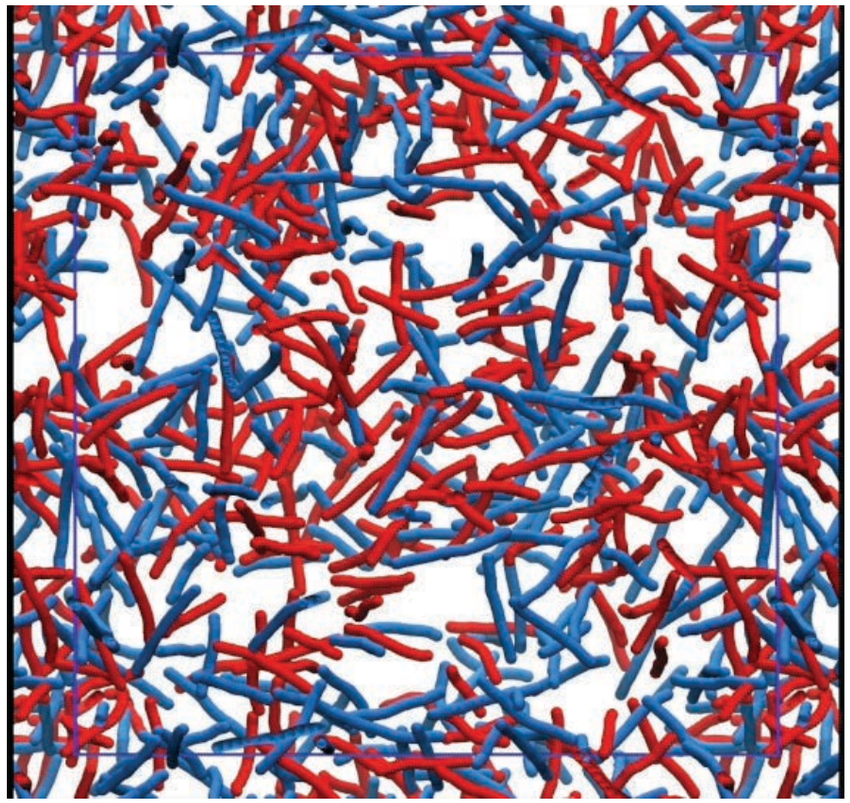
Driving force for phase separation of nuclear DNA?

II







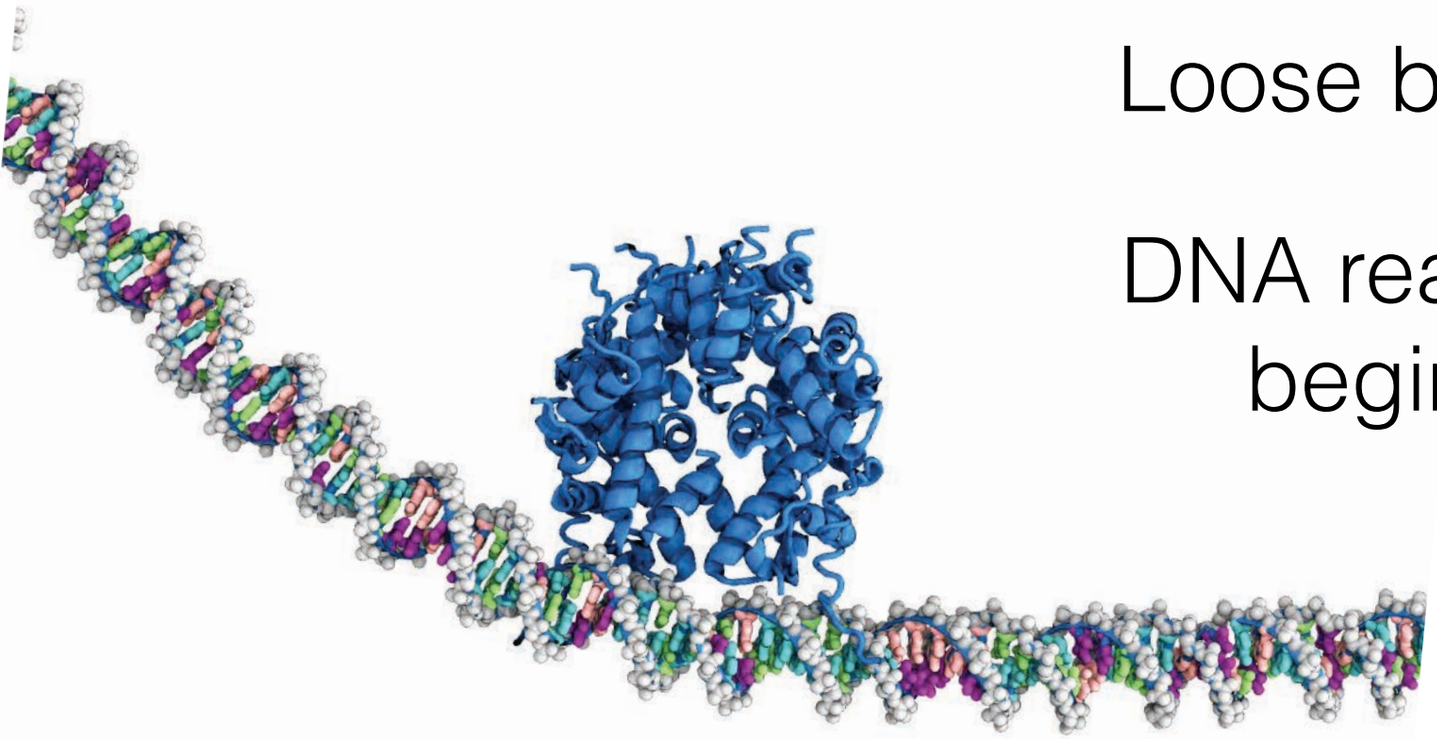


Coarse-grained simulation of a mixture of 250 CG-rich and 250 TA-rich DNA fragments.

The simulation box is 200-nm on edge. The simulation time is 100 microseconds

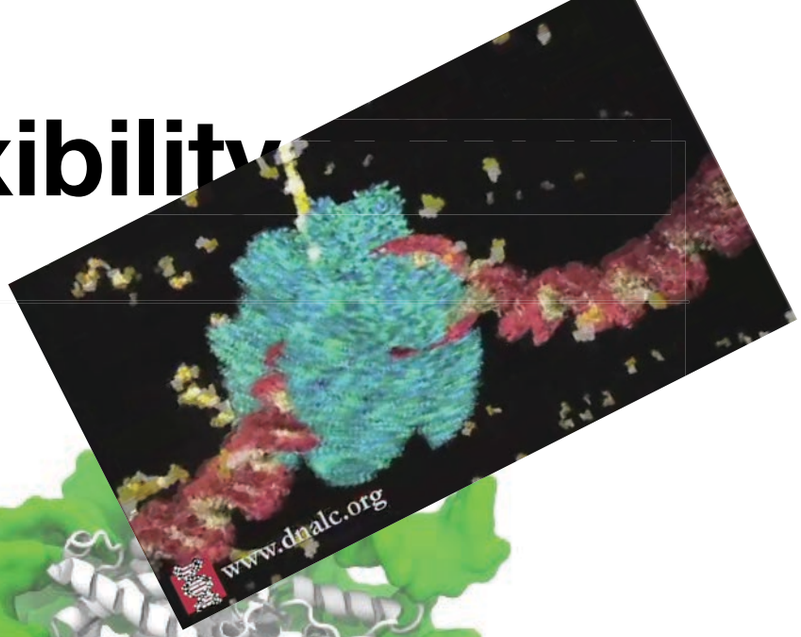
Nucleic Acids Res. **46: 9401 (2018)**

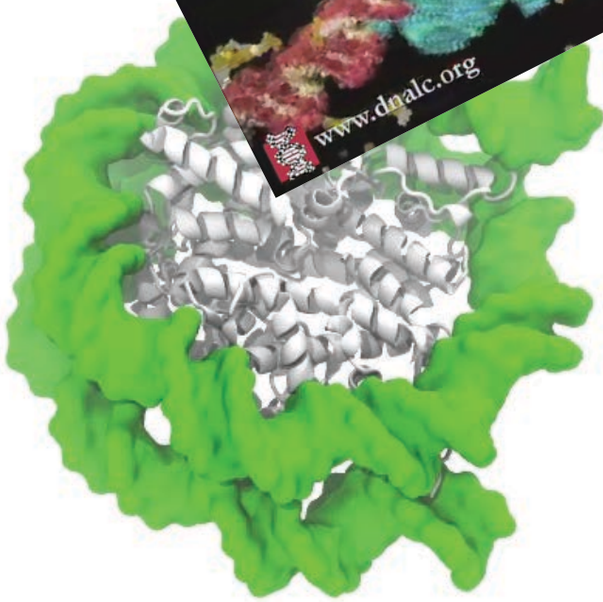
Epigenetic modifications control **flexibility**



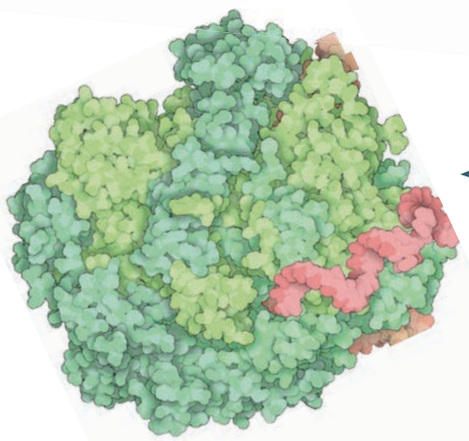
Loose binding:

DNA reading begins





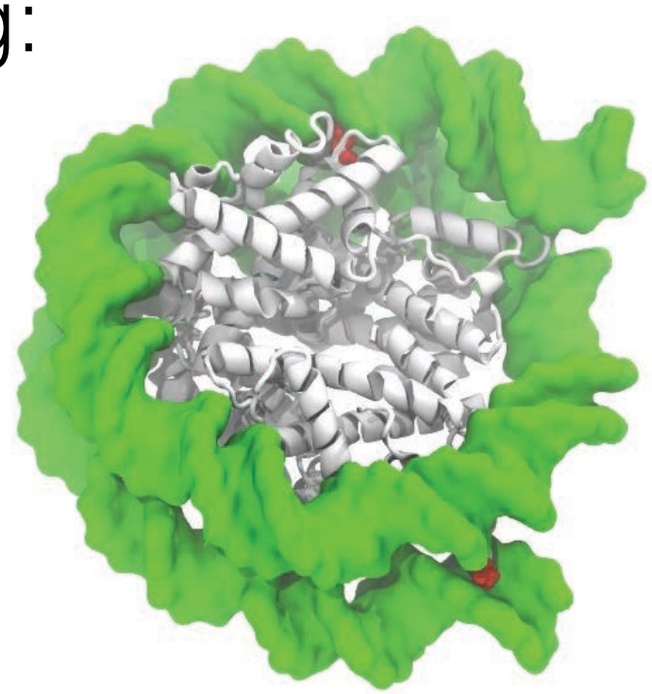
Ngo, et al., *Nature Communications* (2016)



RNA polymerase
reads DNA sequence.

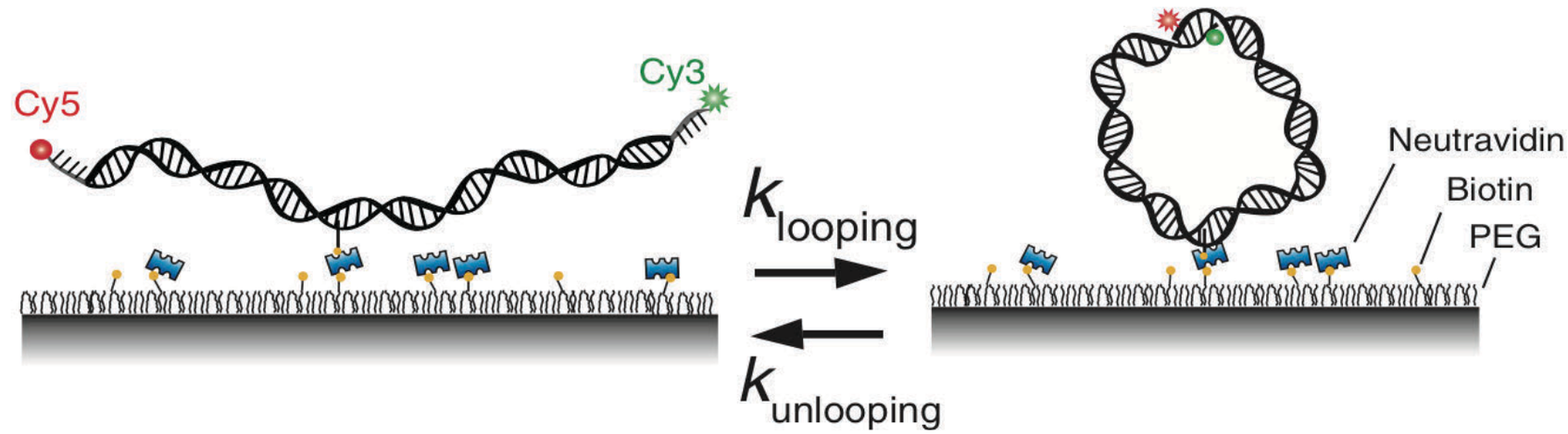
Tight binding:

DNA reading can’t start

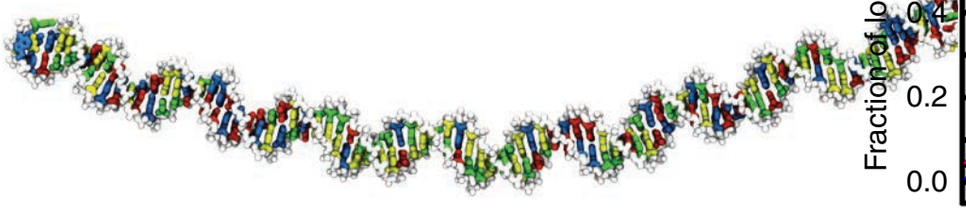


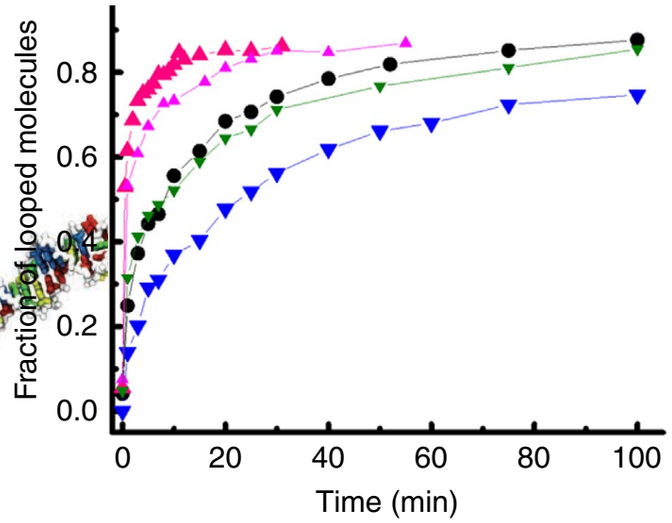
DNA looping assay

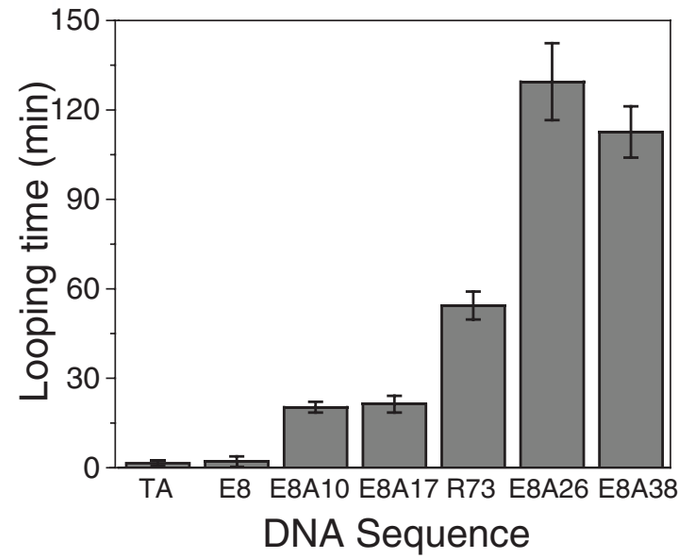
Taekjip Ha



5′- GCTAG TACCTCAATA TAGACTCCCT CCGGTGCCGA GGCCGCTCAA TTGGTCGTAG GACTATCCTC ACCTCCACCG TTTCA







Extremely flexible ⟷ Extremely stiff

Measure average looping time

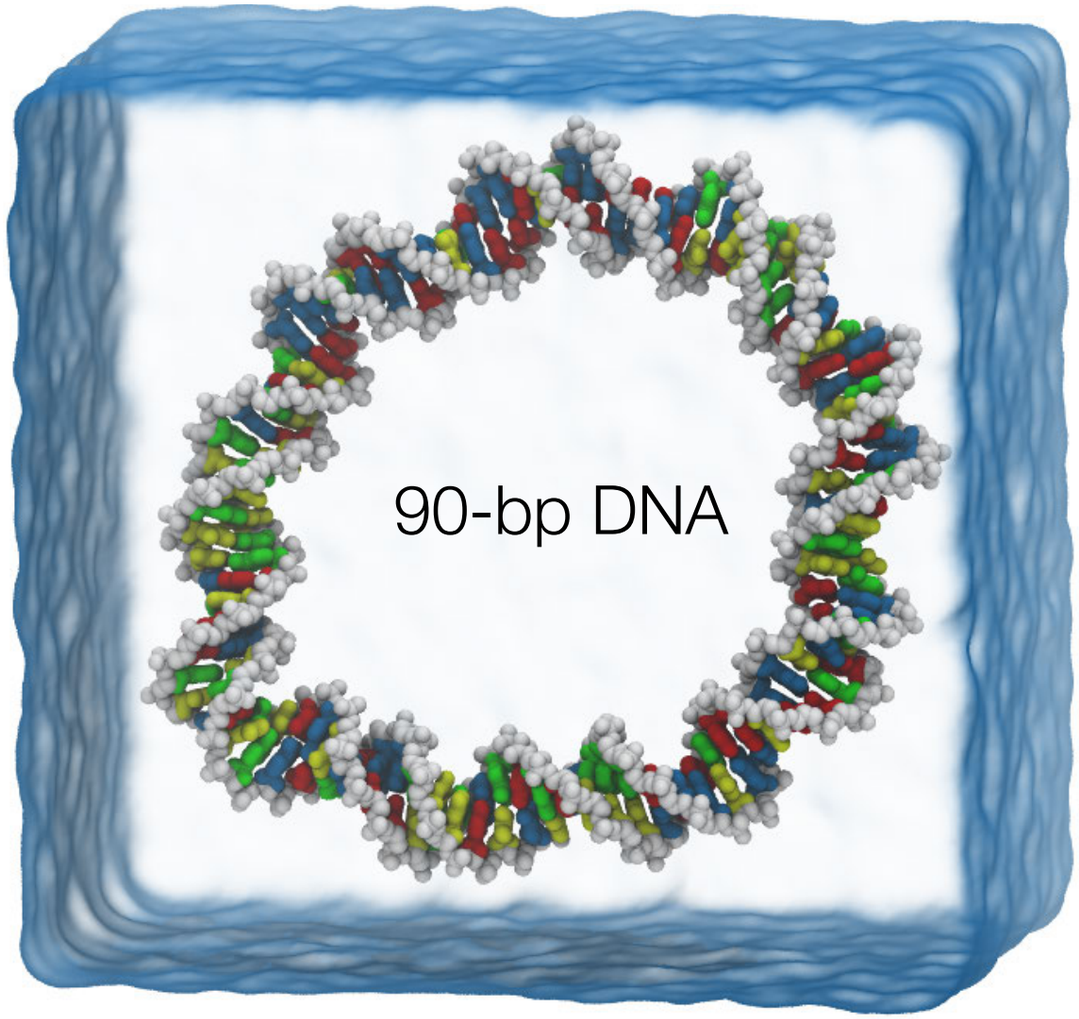
Questions:

- How does DNA sequence program such a broad range of flexibility?

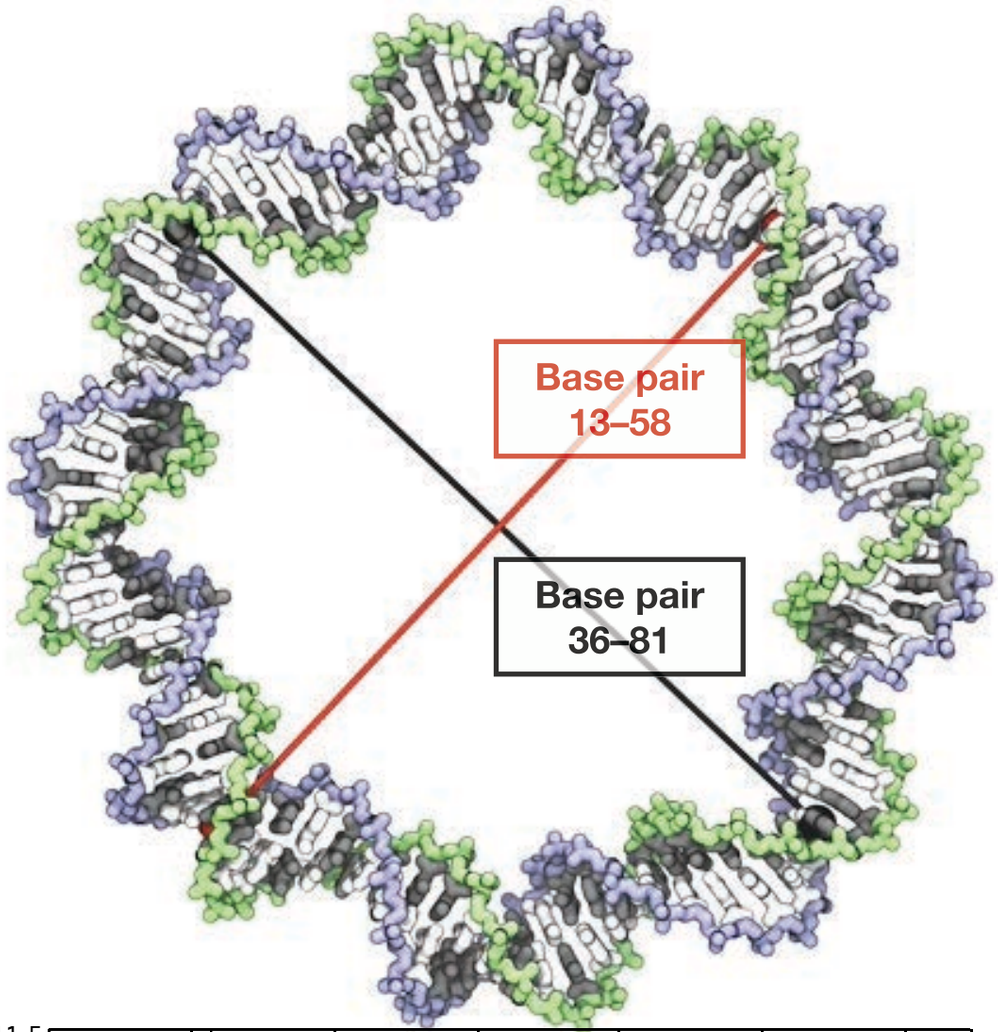
Cloutier and Widom, *Molecular Cell* 2004
Vafabakhsh and Ha, *Science* 2012

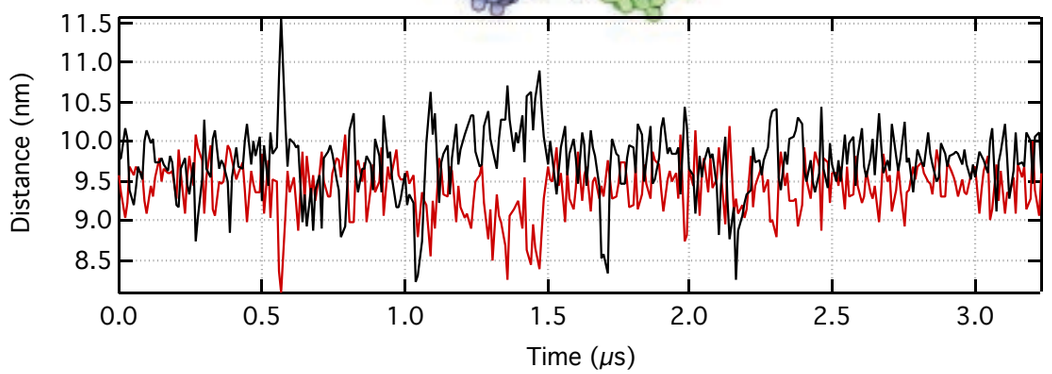
DNA Minicircle as a Model System for Quantifying DNA Flexibility

cagaatccgtgctagtacctcaatatagactccctc**cg**gtgc**cg**a
ggc**cg**ctcaattggt**cg**taggactatcctcacctccaccgtttca



All-atom MD simulations of a DNA minicircle



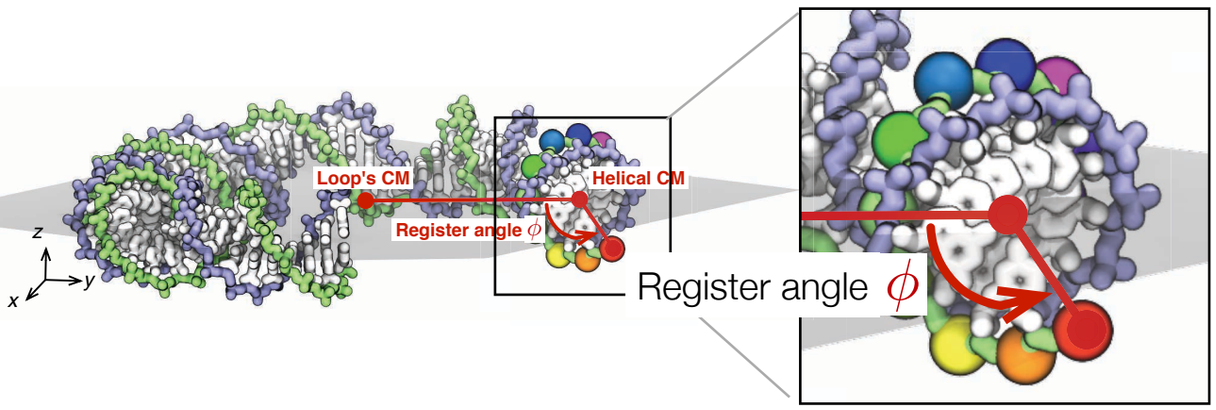


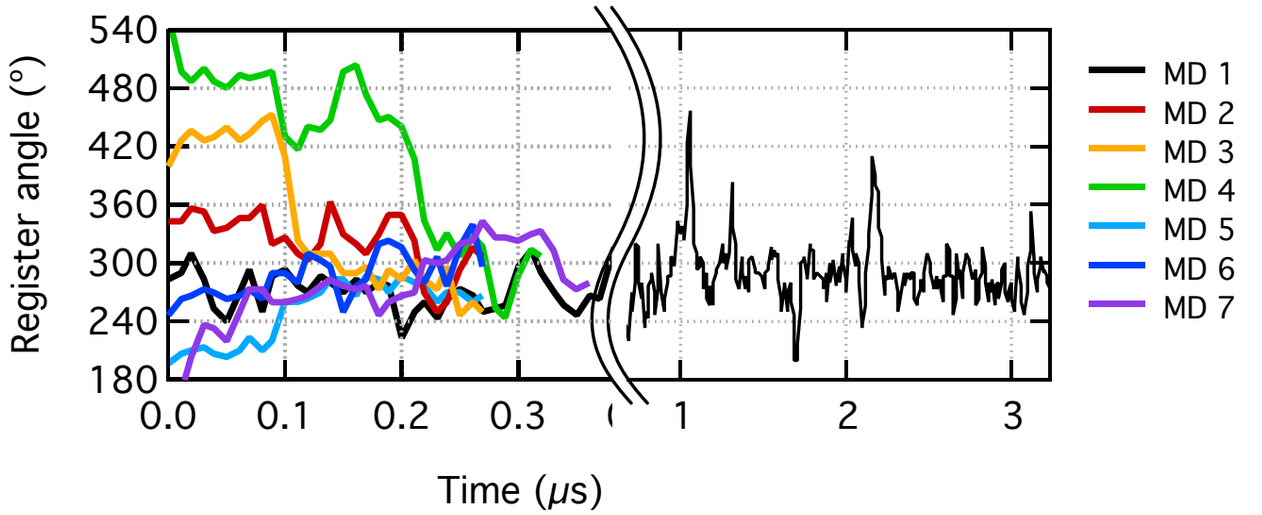
Stable and highly correlated dynamics during 3-μs simulation

Luger et al. ***Nature*** 1997 | Cloutier and Widom, *Molecular Cell* **2004**

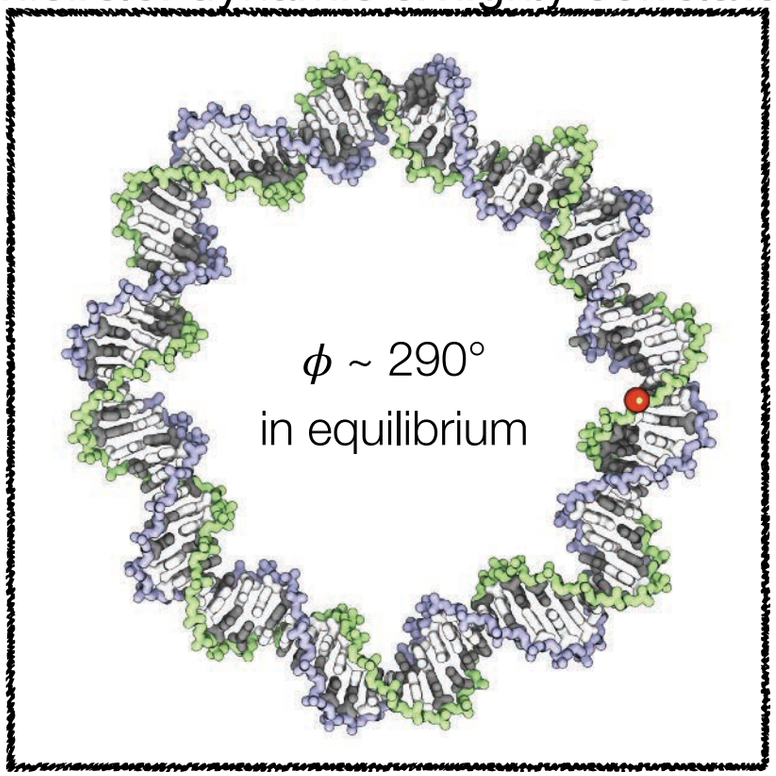
DNA Sequence Programs Preferential Register Angle of Minicircles

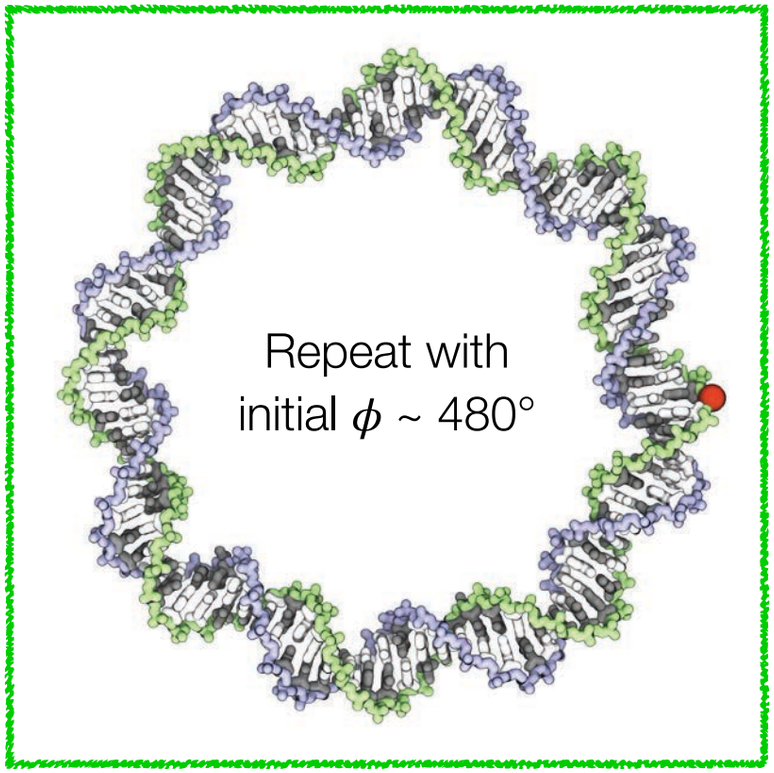
cagaatccgtgctagtacctcaatatagactccctc**cg**gtgc**cg**a
ggc**cg**ctcaattggt**cg**taggactatcctcacctccaccgtttca



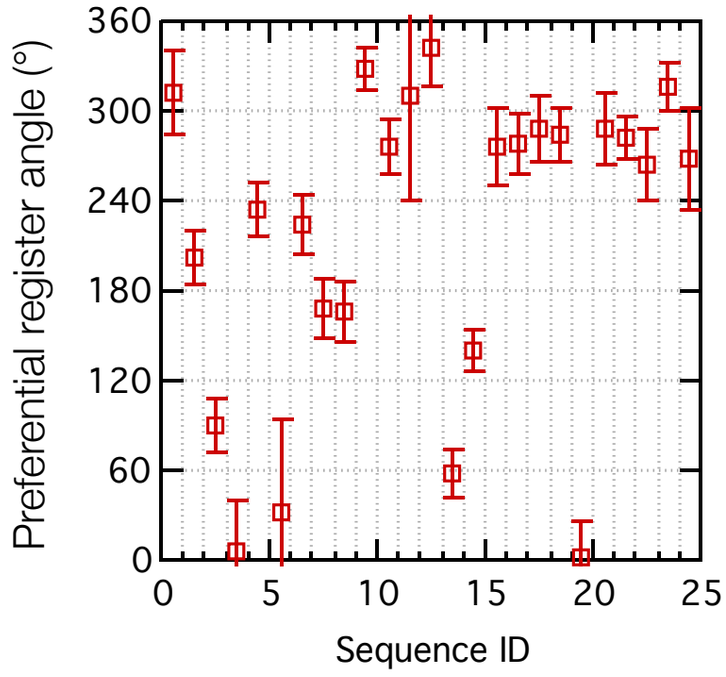


3-μs simulation of 90-bp DNA minicircle: dynamic & highly correlated





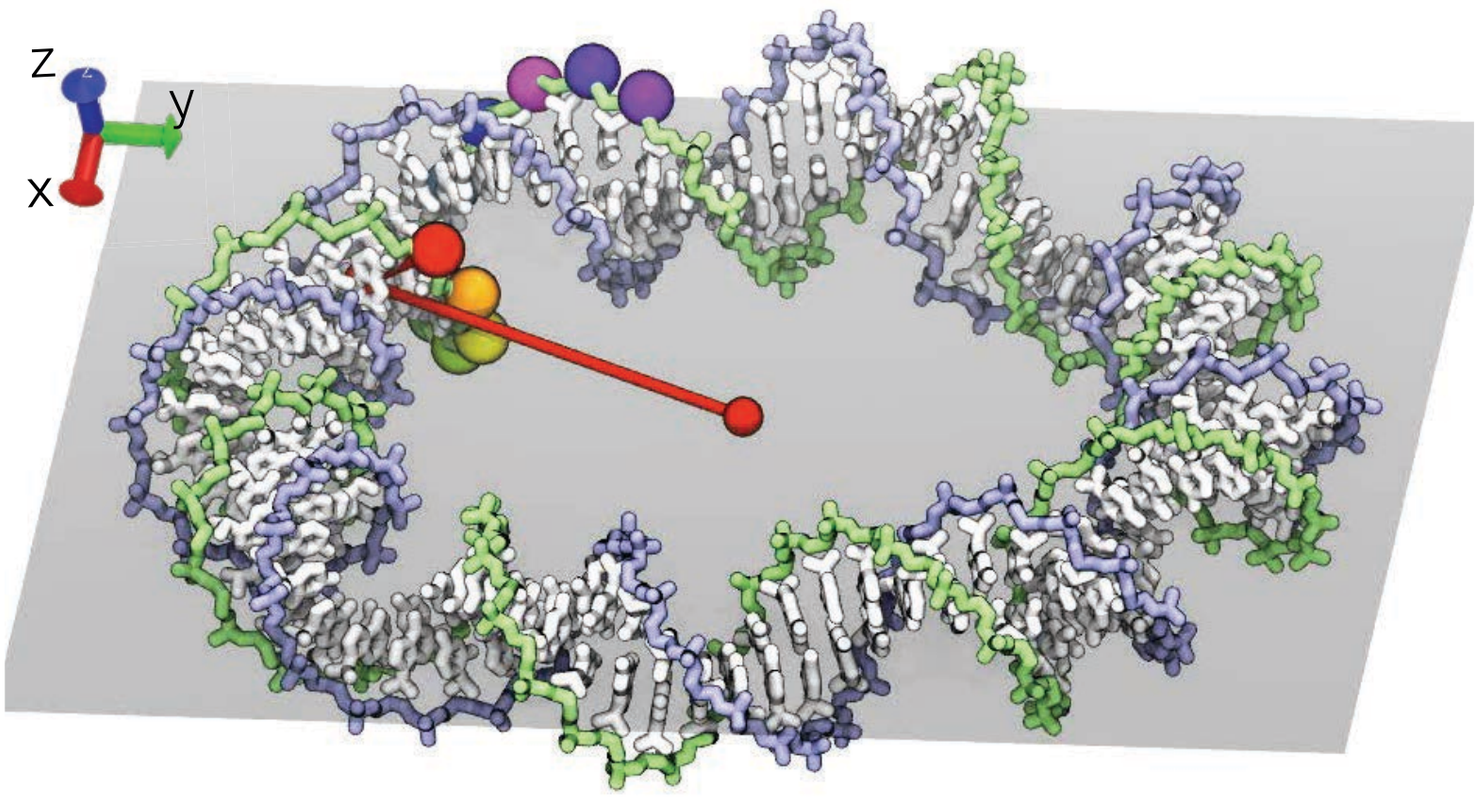
DNA sequence programs the preferential angle.



- The presence of the preferential register angle suggests a non-flat **free energy landscape**.

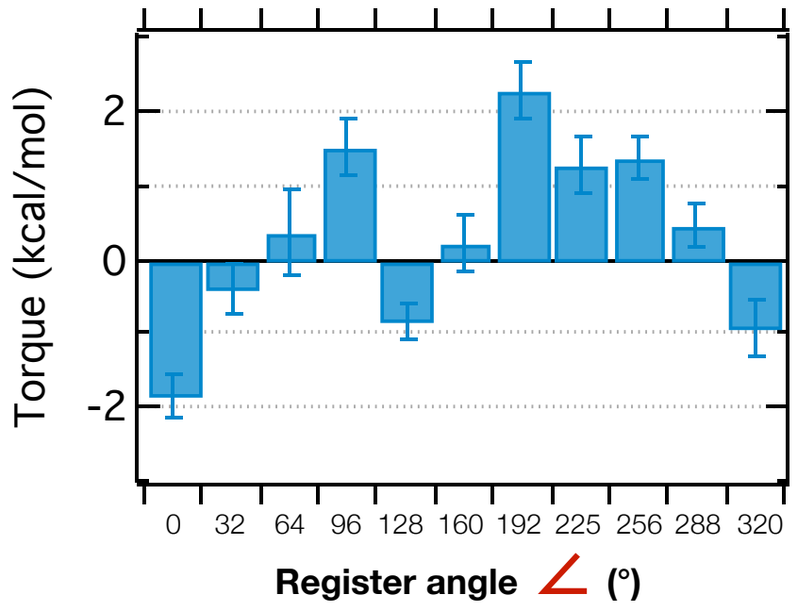
Yoo et al. to be published

Preferential Register Angle Is the Global Minimum of Free Energy Landscape

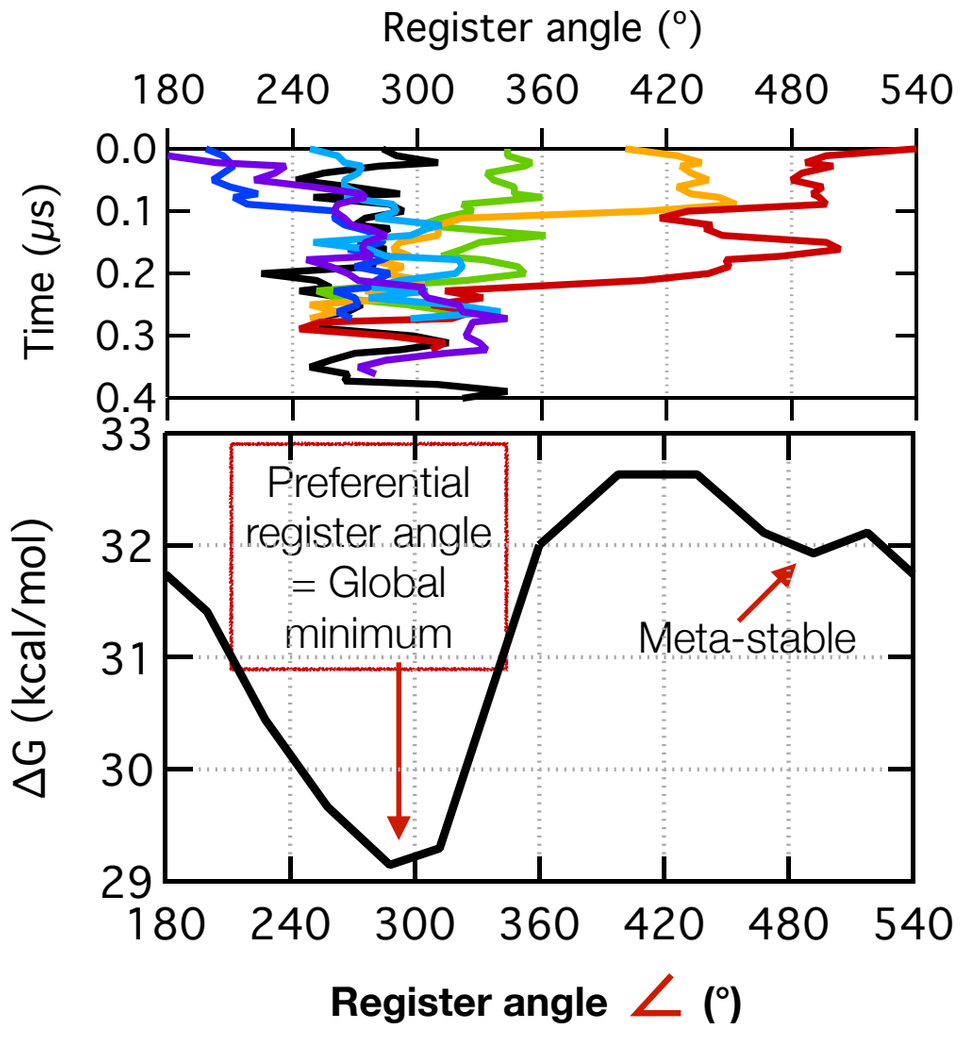


- We change the register angle ∠, from 0° to 360° with 32° increment.
- In this movie, rotation is accelerated for visualization purpose.
- At a given angle, we compute the average torque for ~0.5 μs.
- In total, 5.5 μs simulation.

Compute torque

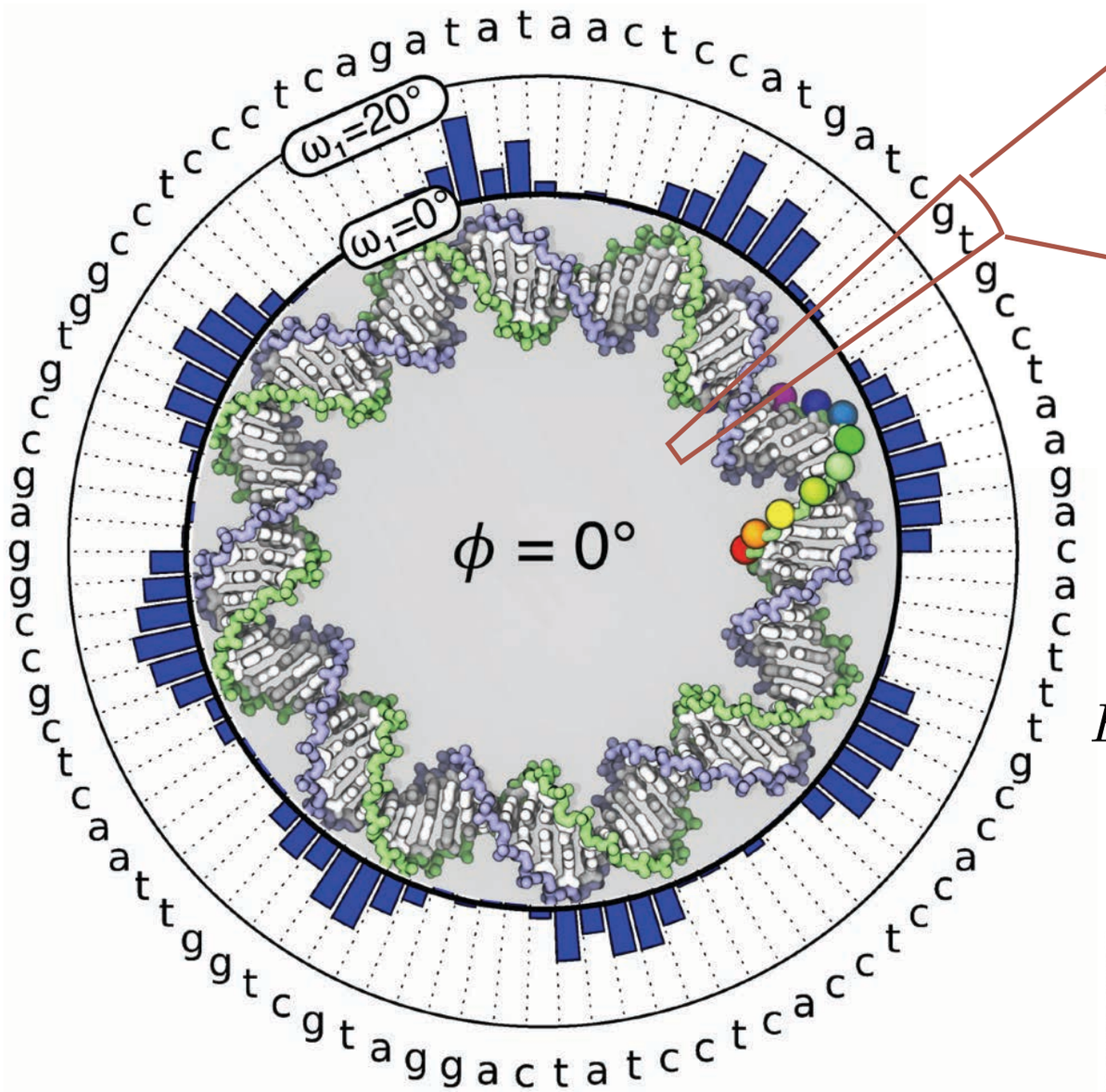


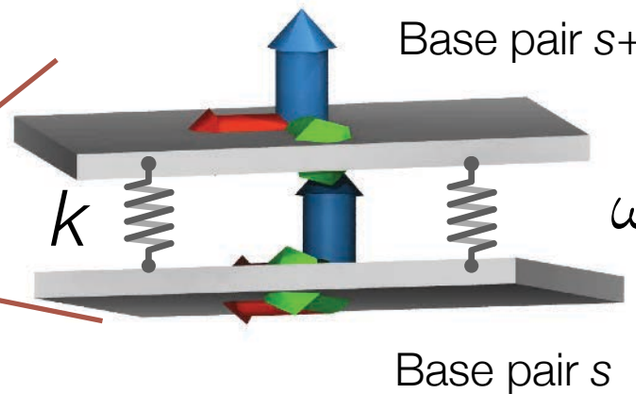
Integrate torque



Yoo et al. to be published

Why Is Free Energy Landscape Non-Flat?





$\omega_{1,s}(\phi) =$ Degree of bending of dinucleotide step s at a given register angle

Bending of dinucleotide step is **non-uniform** along the circle, $s = 1 \cdots 90$.

$$E(\phi) = \sum_{s=1}^{90} k(\text{dinucleotide type of } s, s+1)\, \omega_{1,s}(\phi)^2$$

Sequence-dependent **dinucleotide bending stiffness**

- To minimize the free energy, avoid bending stiff dinucleotides: mechanical minimal frustration.

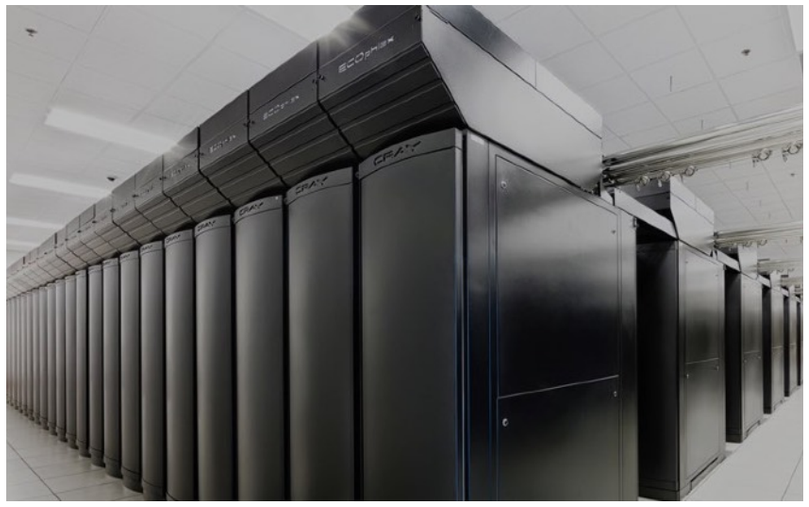
- Then, which dinucleotide steps are stiff and which are flexible?

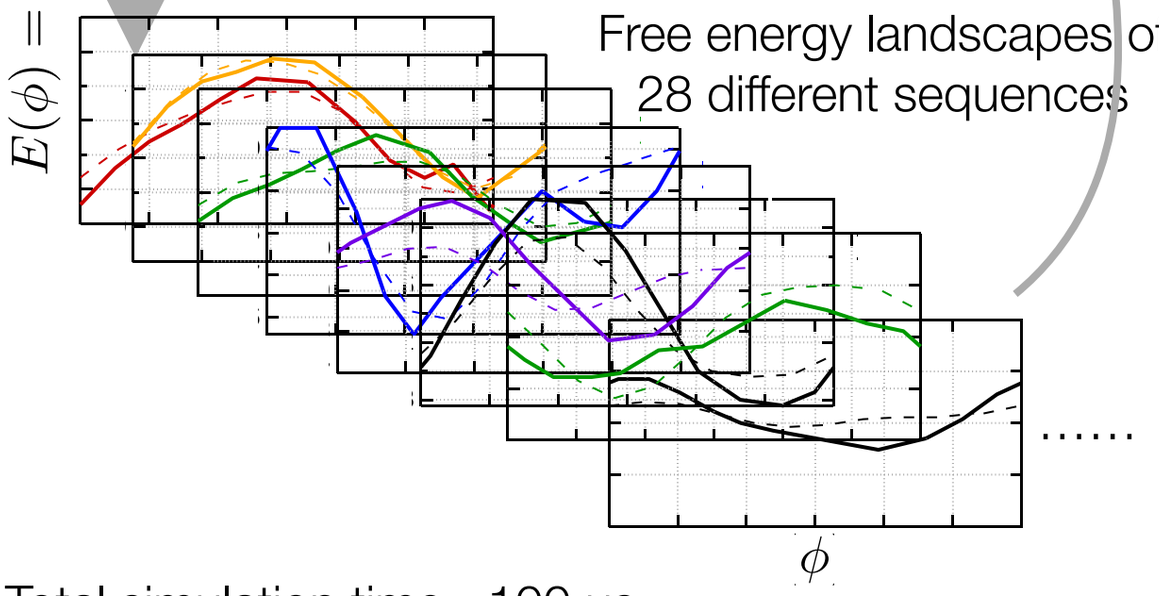
Yoo et al. to be published

Constructing Dinucleotide Bending Stiffness Matrix Through High-Throughput MD Simulations of DNA Minicircles

Automated high-throughput computation of free energy landscapes

Blue Waters Supercomputer





Total simulation time ~100 μs

$$E(\phi) = \sum_{s=1}^{90} k(\text{dinucleotide type of } s, s+1)\omega_{1,s}(\phi)^2$$

Optimize **Dinucleotide Bending Stiffness Matrix** using the free energy landscapes.

Second base of dinucleotide step

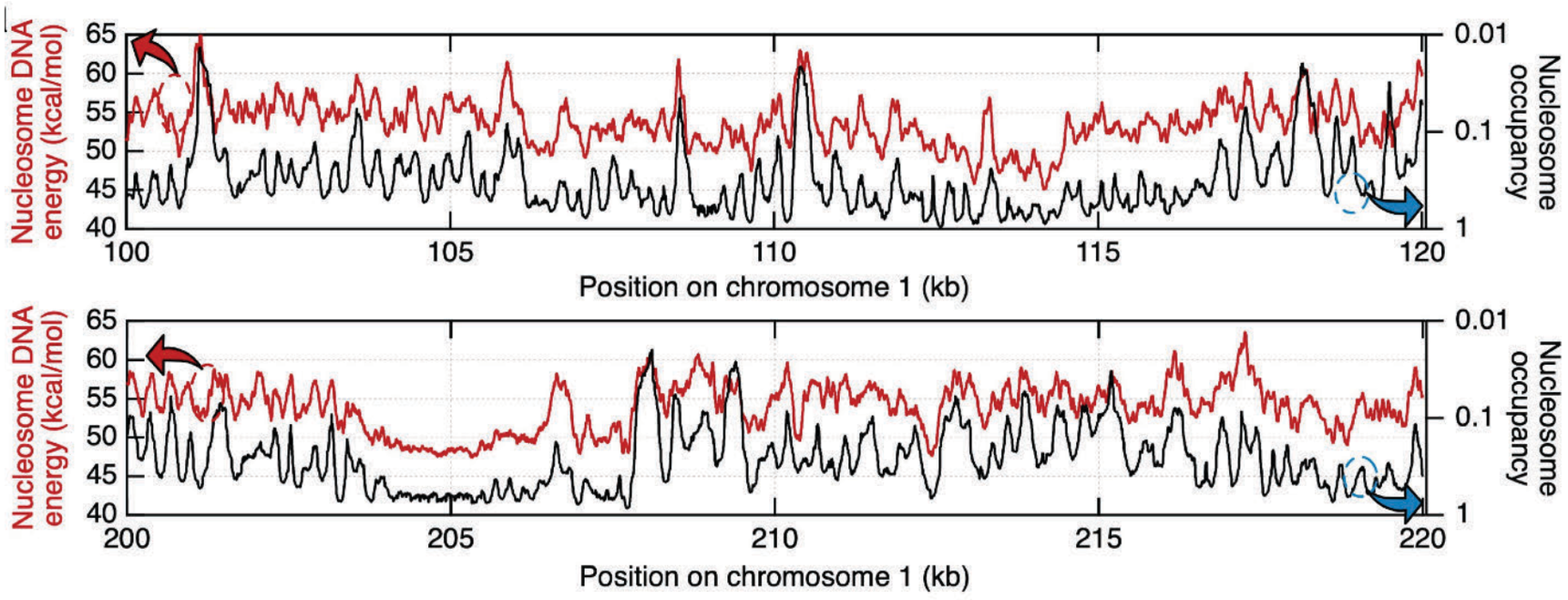
First base of dinucleotide step	A	G	mG	T	C	mC
T	12	6	–	68	60	120
C	6	12	–	35	34	47
mC	–	–	30	–	–	–
A	68	35	–	48	44	39
G	60	34	–	44	58	50
mG	120	47	–	39	50	38

Stiffness (kcal/mol/rad²): >110, 90, 70, 50, 30, <10

mG: G of mCpG sites

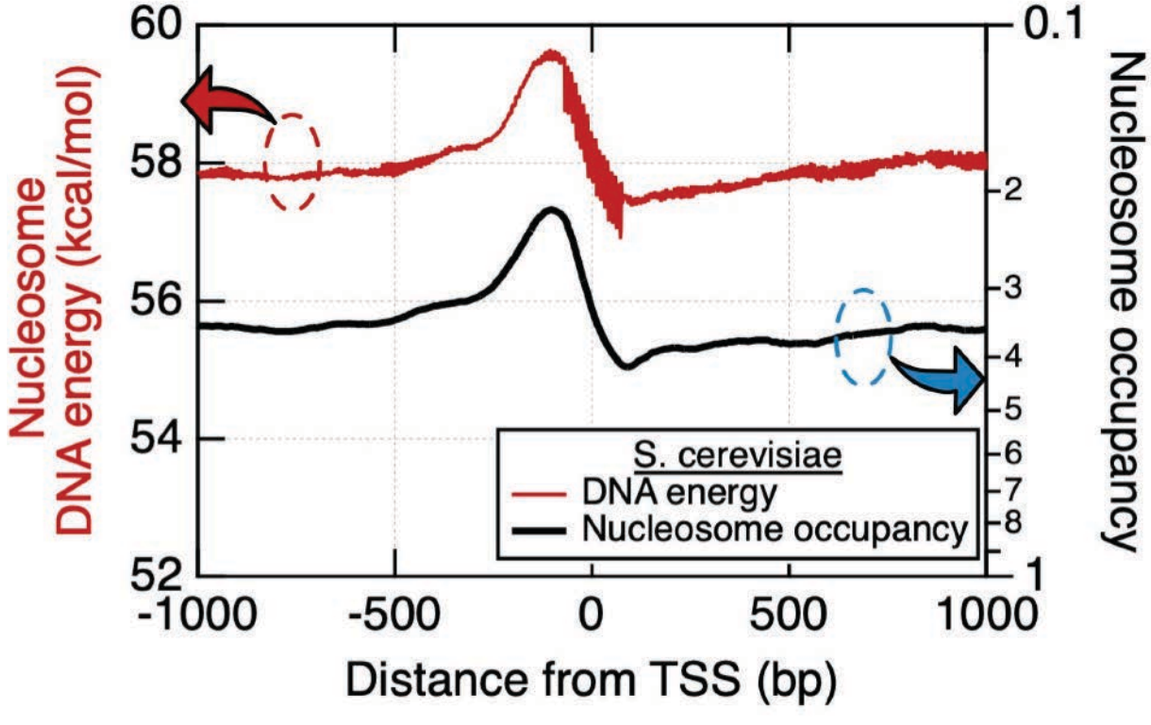
Yoo et al. to be published

Connecting Physics to “Omics”



Bending energy versus experimental nucleosome occupancy two 20-kb segments of chromosome 1 of *Saccharomyces cerevisiae.*

Nucleosome occupancy and bending energy averaged over all transcription start sites (TSS) in the genome of Saccharomyces cerevisiae.



Jejoong Yoo

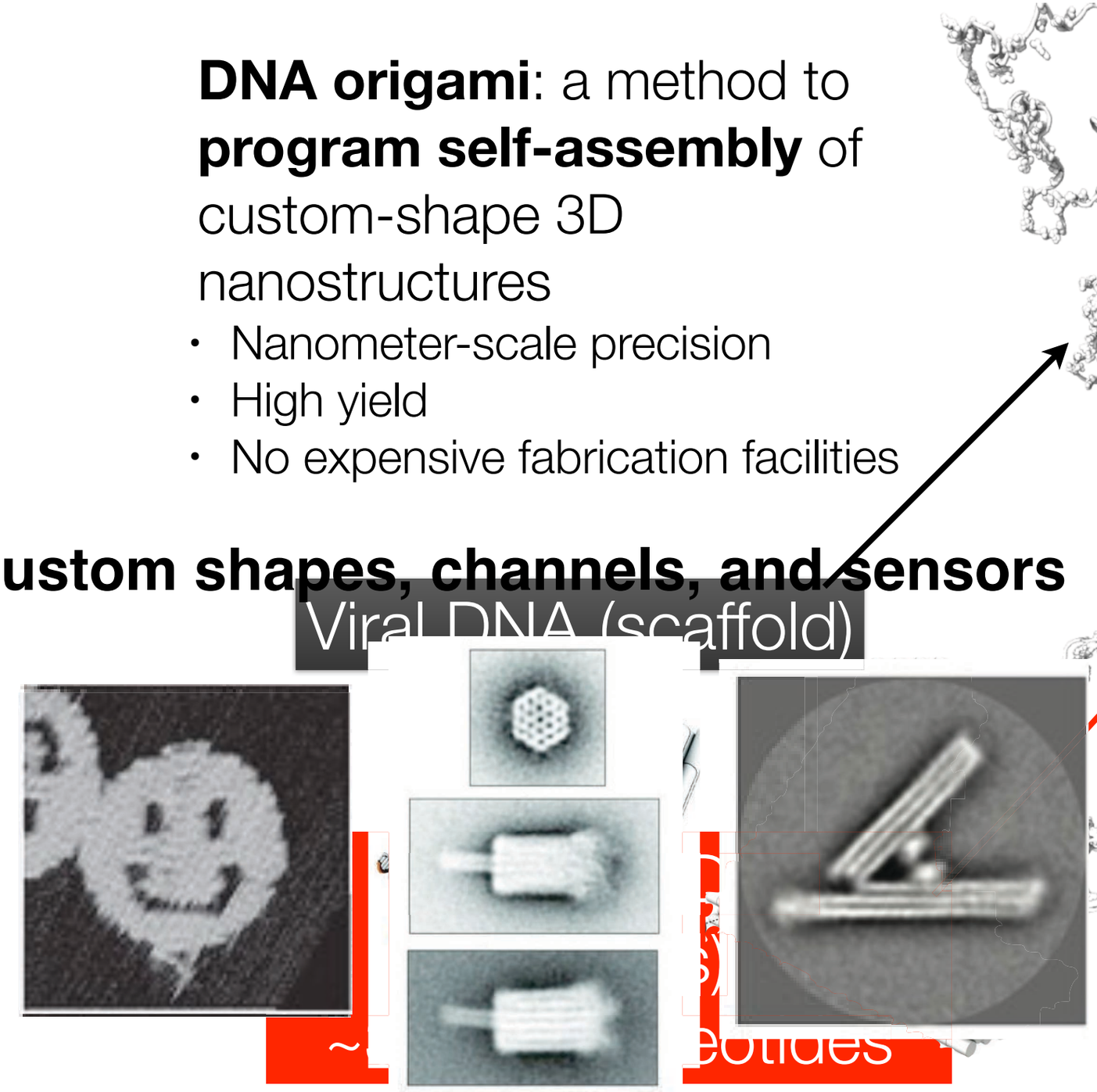
Yoo et al. to be published

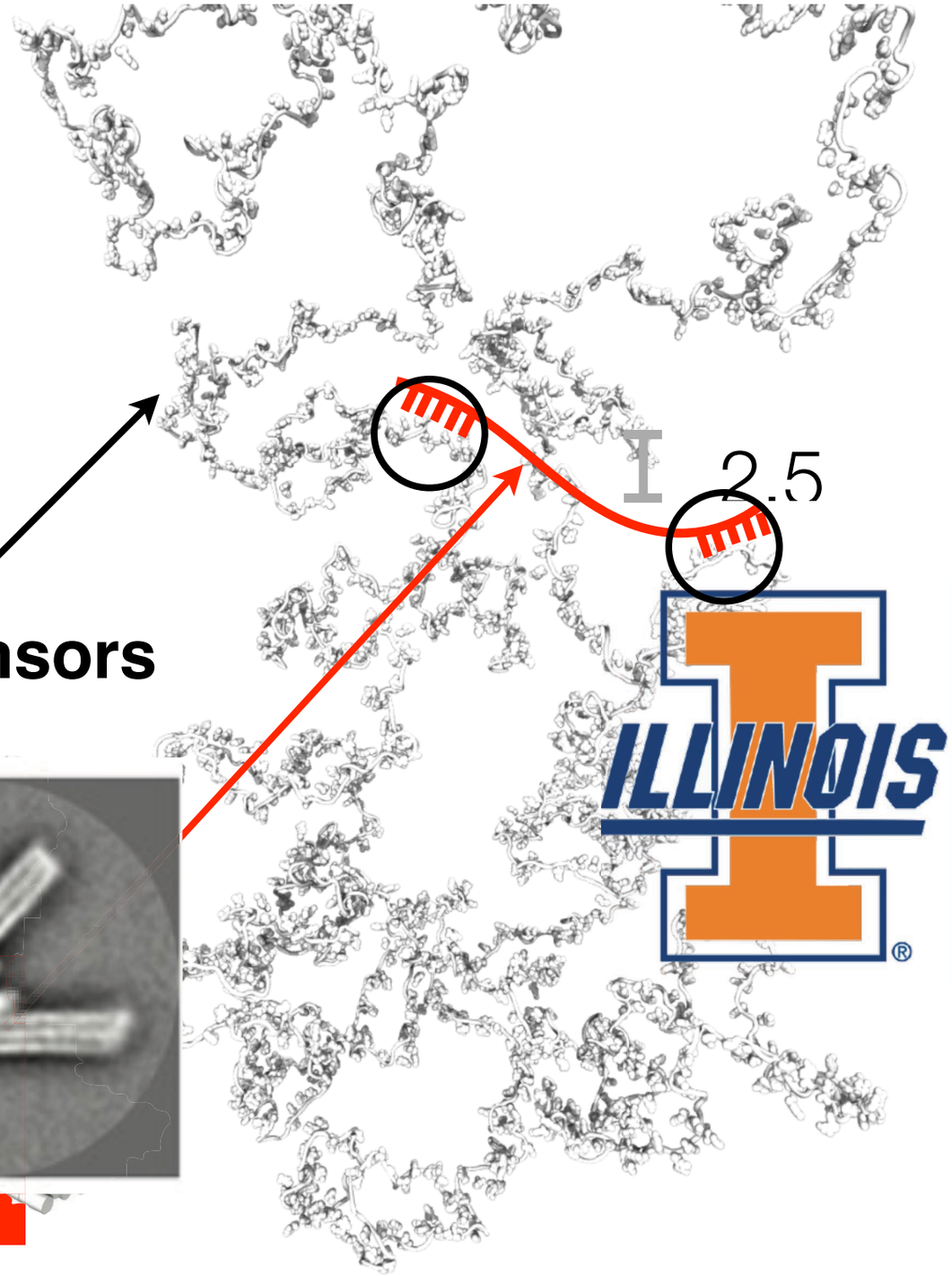
DNA, a building material

DNA origami: a method to **program self-assembly** of custom-shape 3D nanostructures

- Nanometer-scale precision
- High yield
- No expensive fabrication facilities

Custom shapes, channels, and sensors

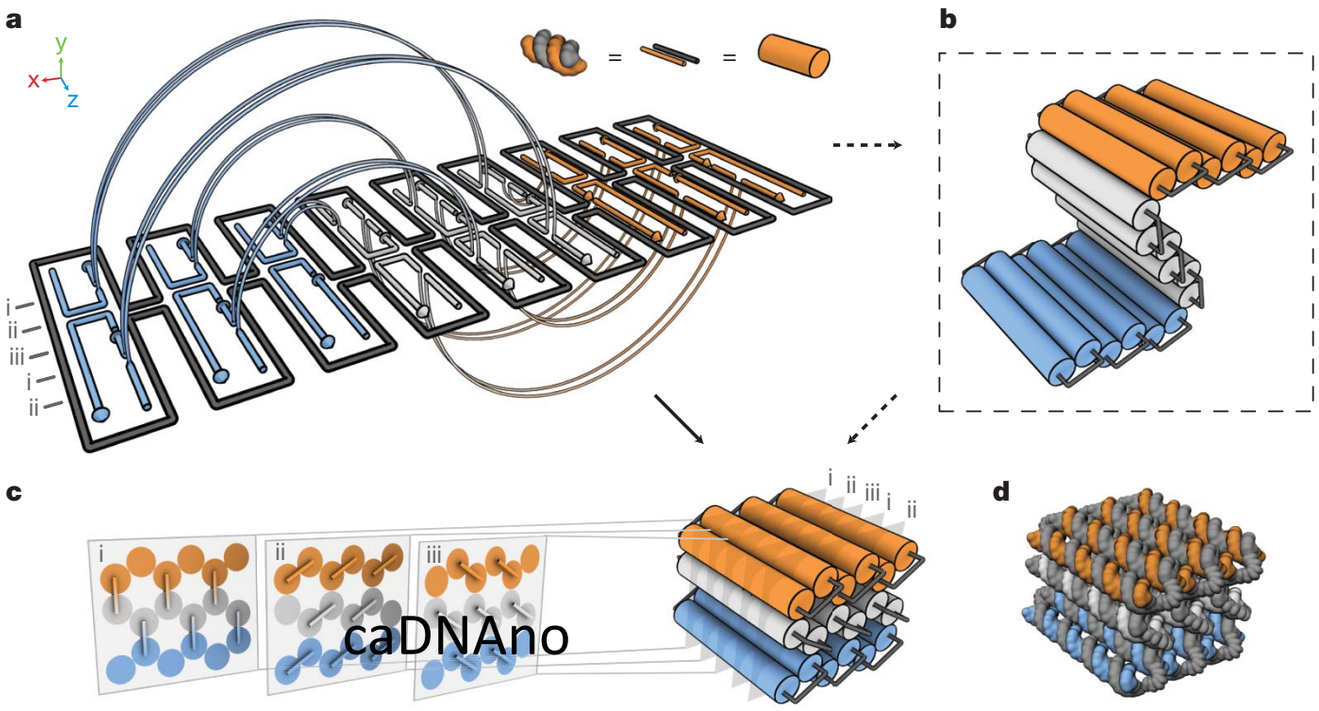




Nadrian Seeman
Paul Rothemund
William Shih
Hendrik Dietz

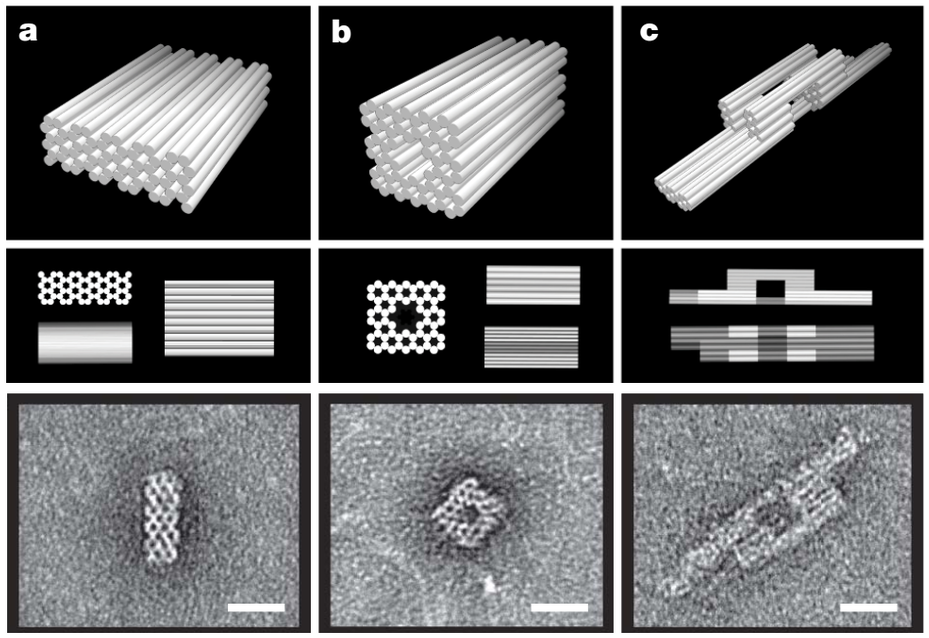
For illustration, an unfolding trajectory at a high temperature is played backward.

Design and characterization of DNA nanostructures

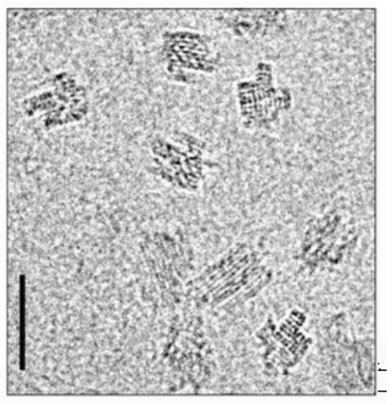


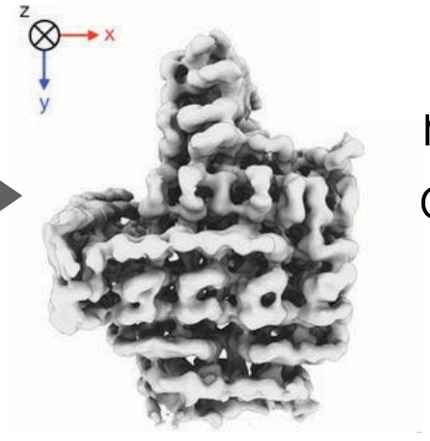
S.M. Douglas, at el. Nature (2009)

Computer-aided design of DNA origami with caDNAno (Shih group, Harvard U.)



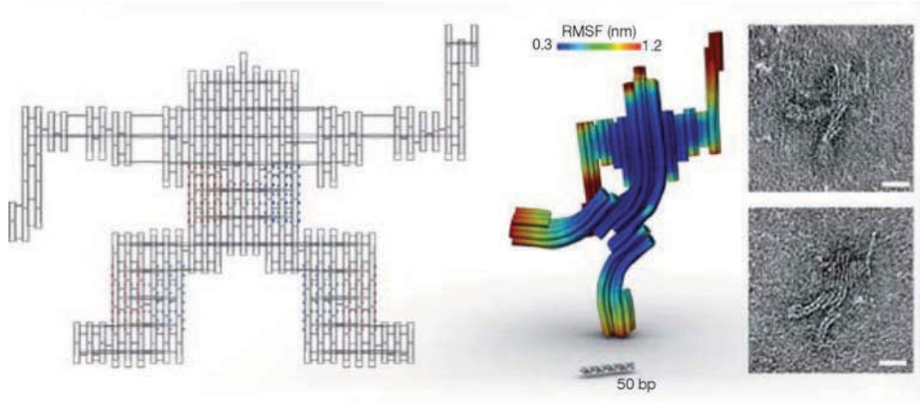
Transmission electron microscopy and/or atomic force microscopy validates the design





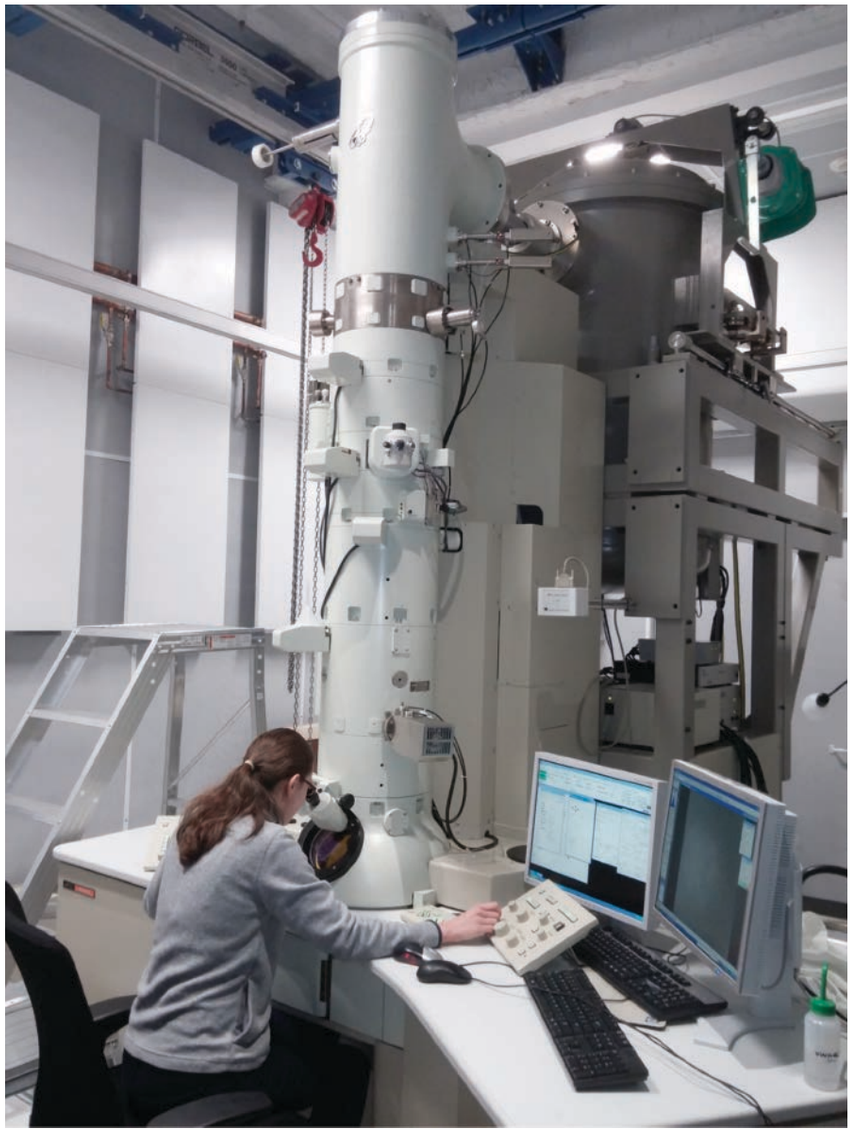
Cryo-EM reconstruction, the only experimentally derived structural model

Bai, … , Dietz, PNAS (2012)



CanDo (Mark Bathe, MIT)

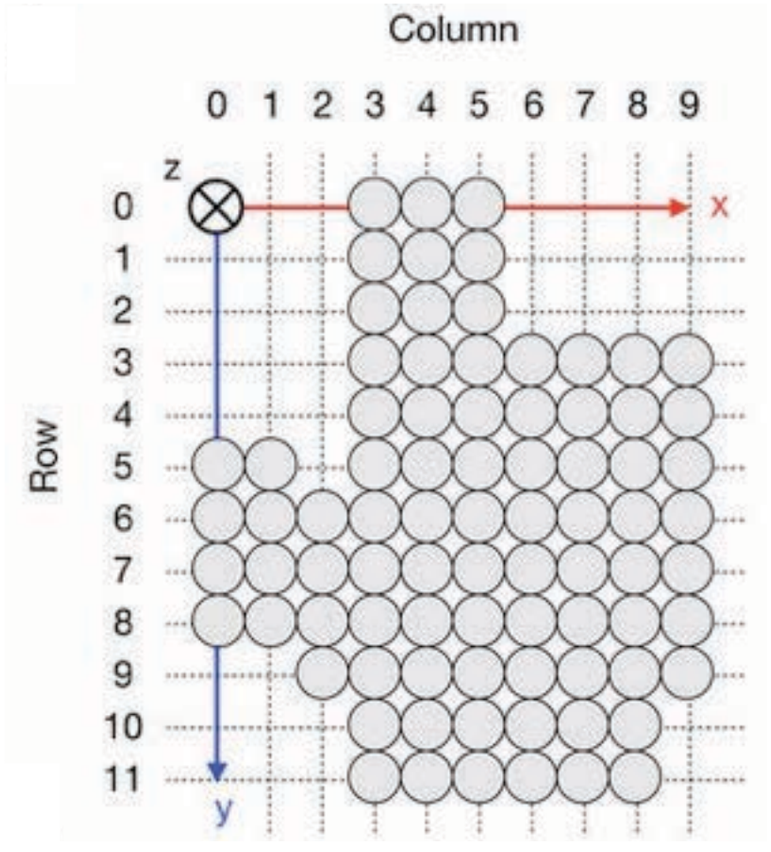
Cryo-EM reconstruction versus all-atom simulation

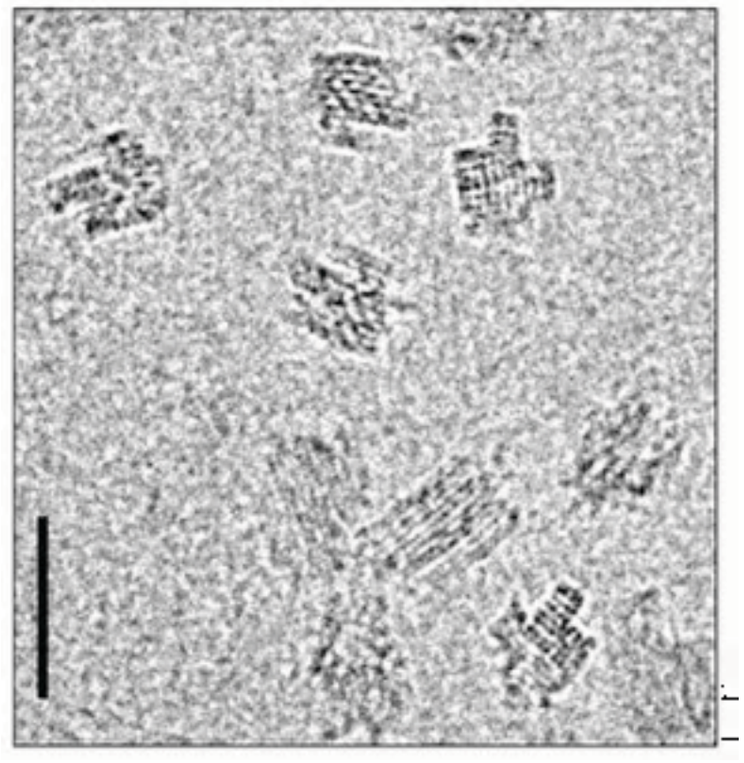


High-resolution cryo-electron microscope

Petascale computer system

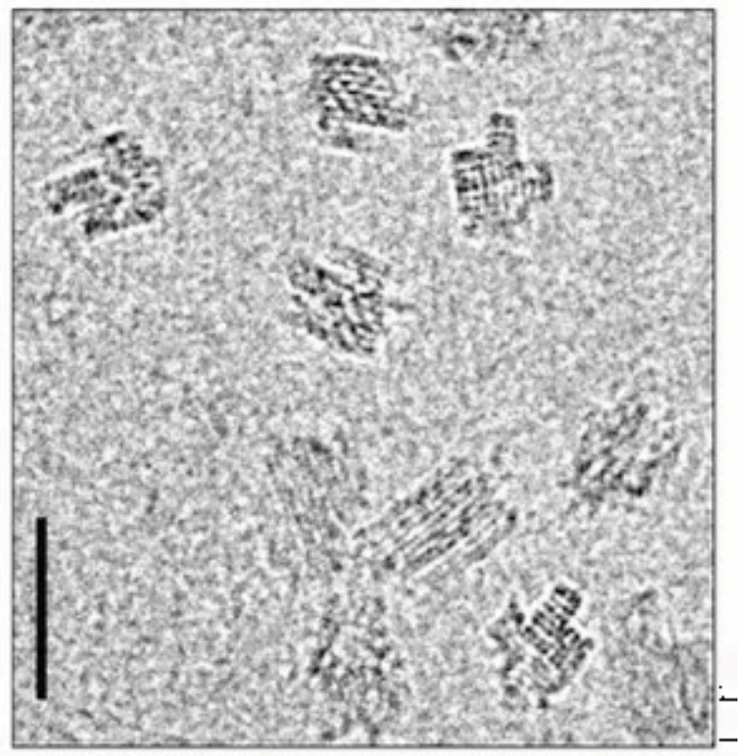
Cryo-EM reconstruction versus all-atom simulation

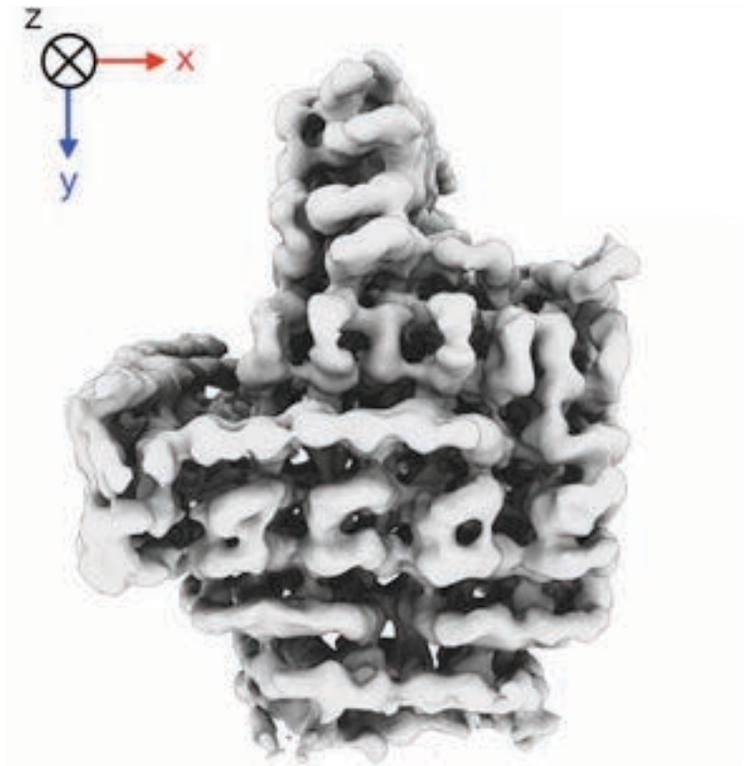




Bai *et al*, PNAS 109:20012 (2012)

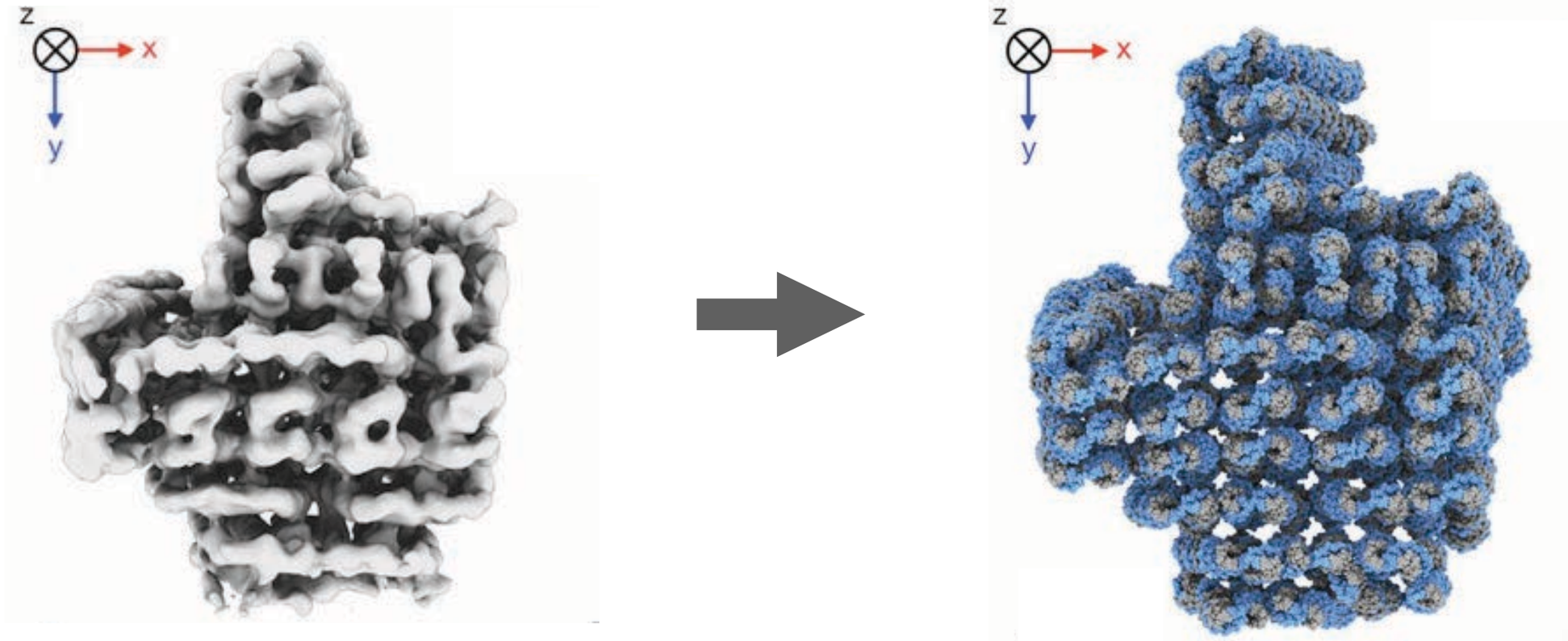
Cryo-EM reconstruction versus all-atom simulation





Bai *et al*, PNAS 109:20012 (2012)

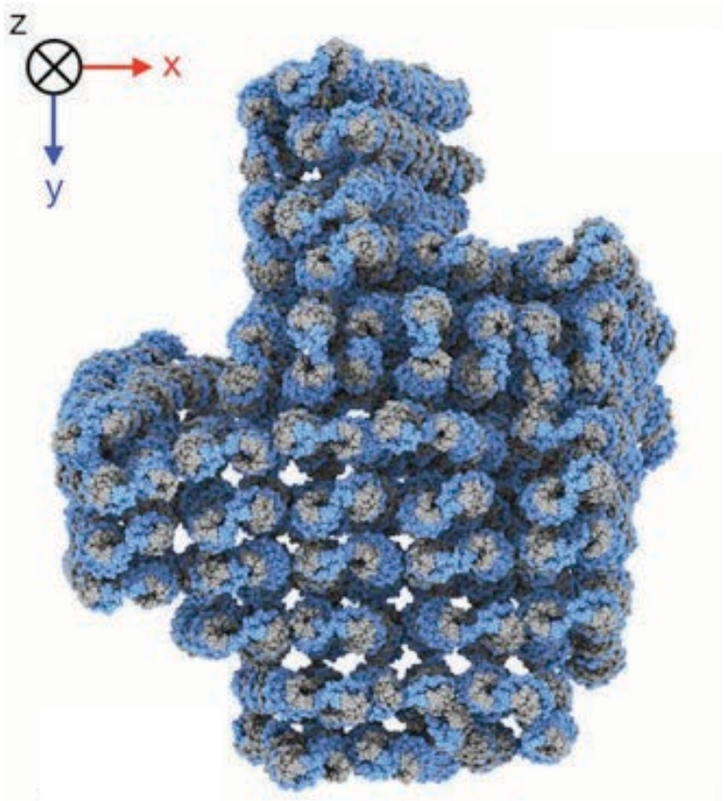
Cryo-EM reconstruction versus all-atom simulation



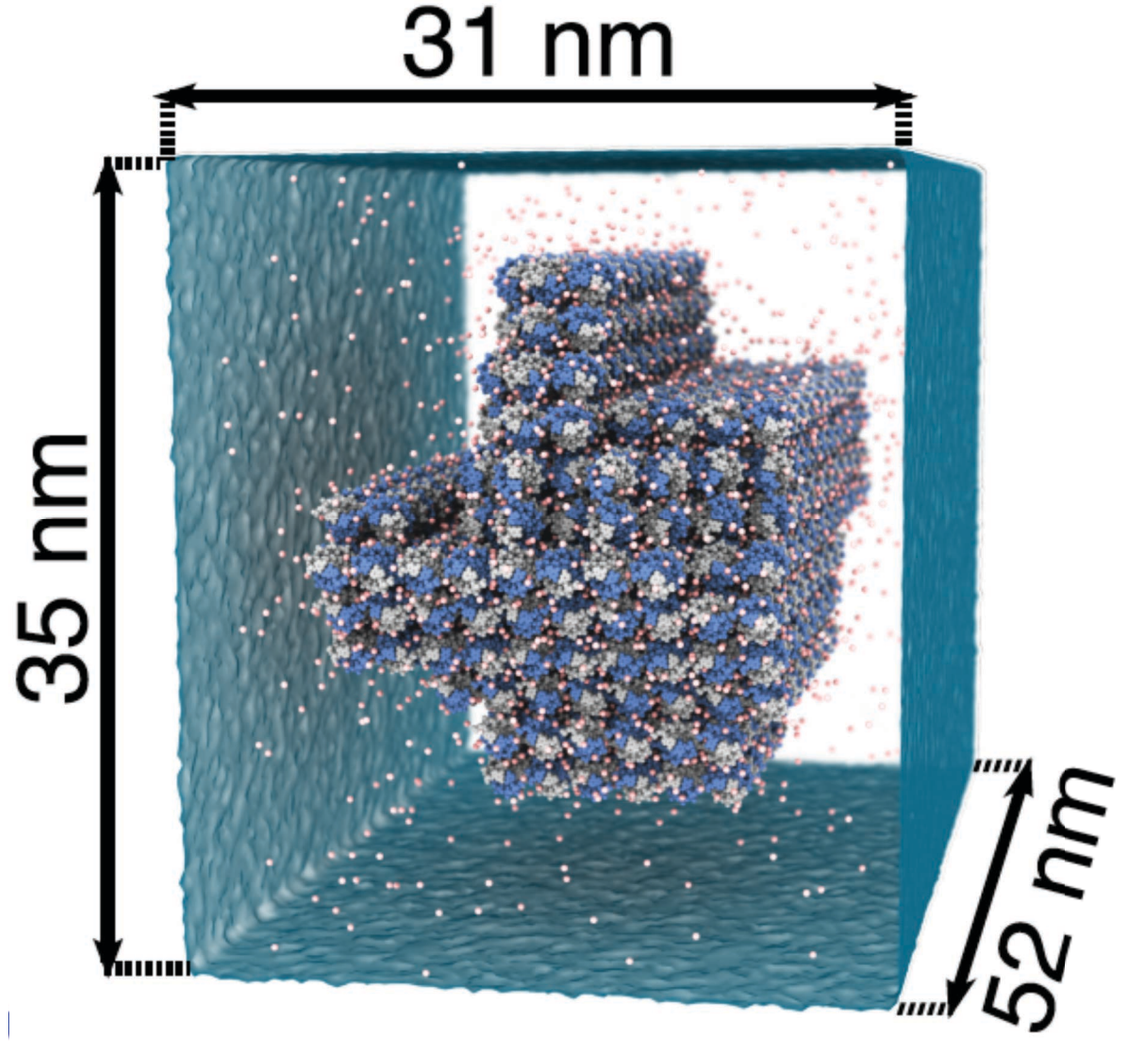
Pseudo-atomic model

Bai *et al*, PNAS 109:20012 (2012)

MD simulation of the cryo-EM object starting from a caDNAno design

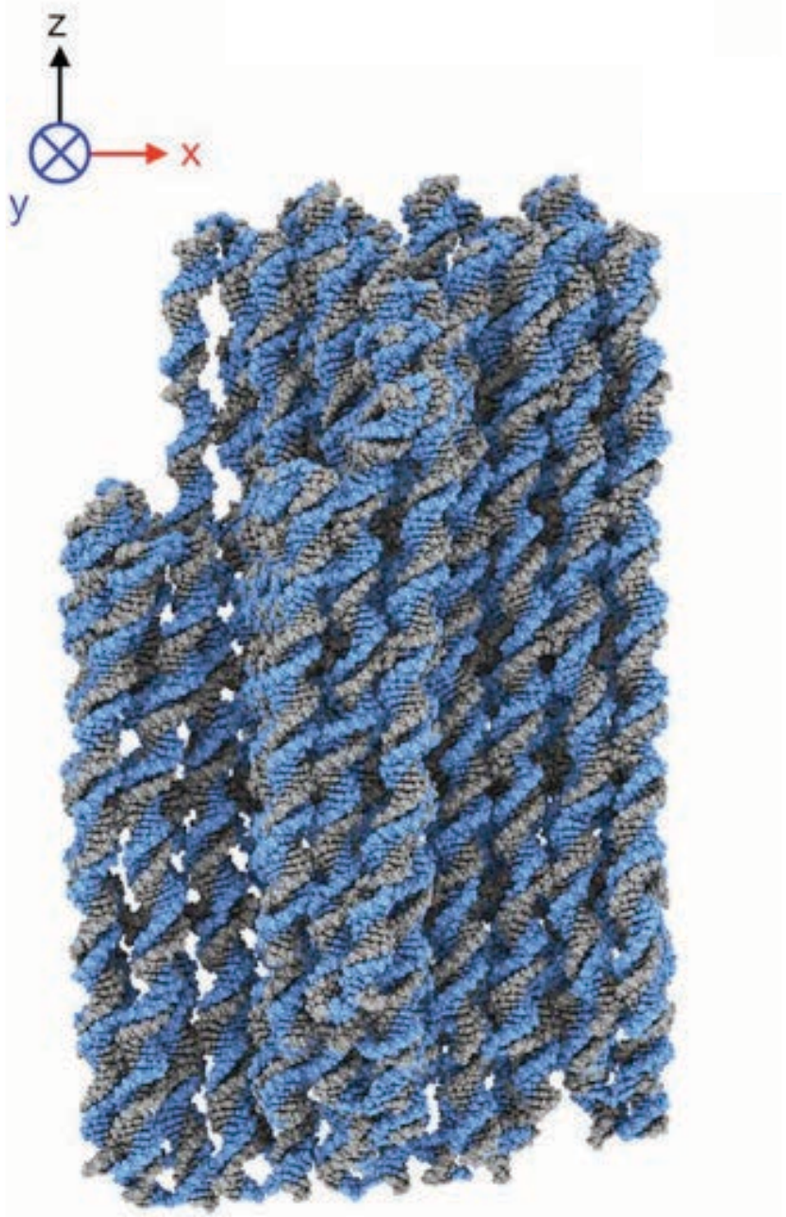


Bai *et al*, PNAS 109:20012 (2012)

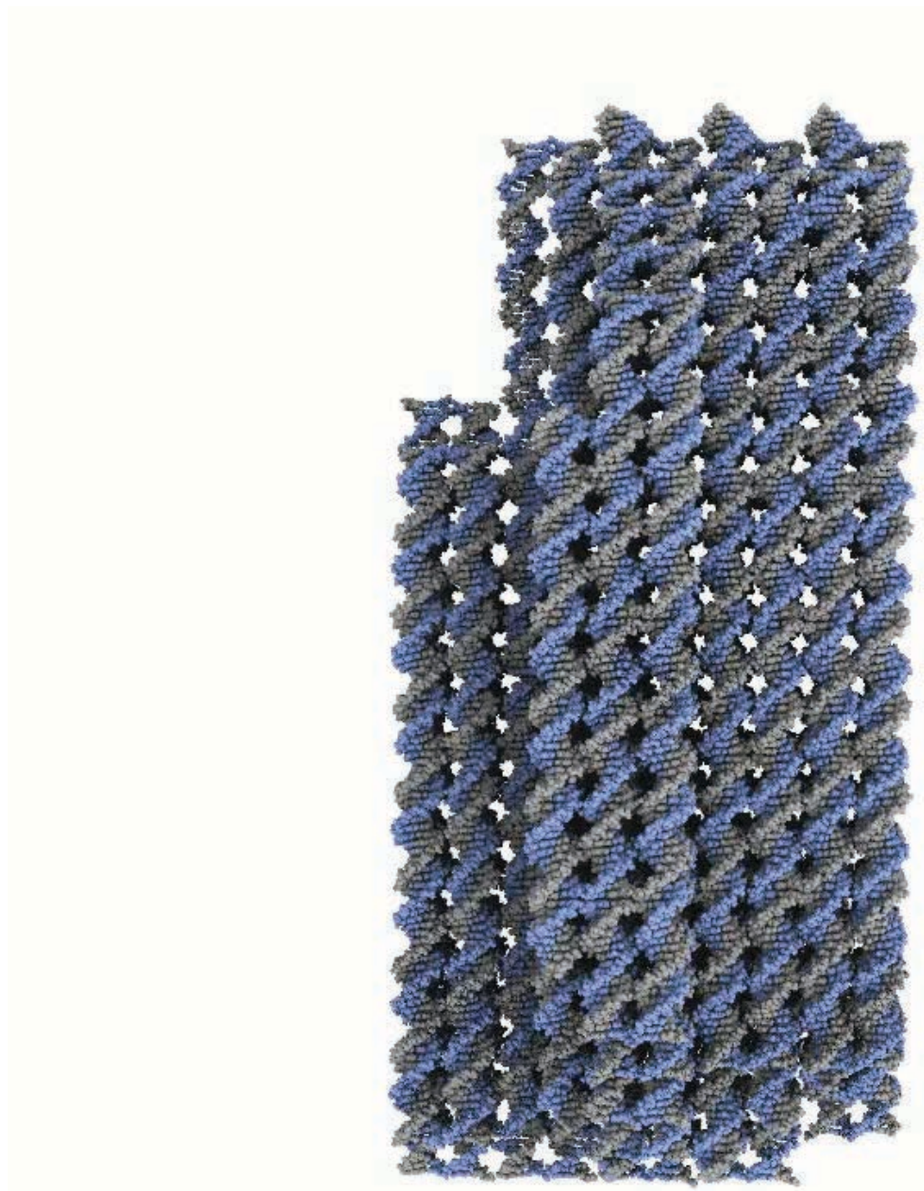


7M atom solvated model
~200 ns MD trajectory

MD simulation of the cryo-EM object starting from a caDNAno design

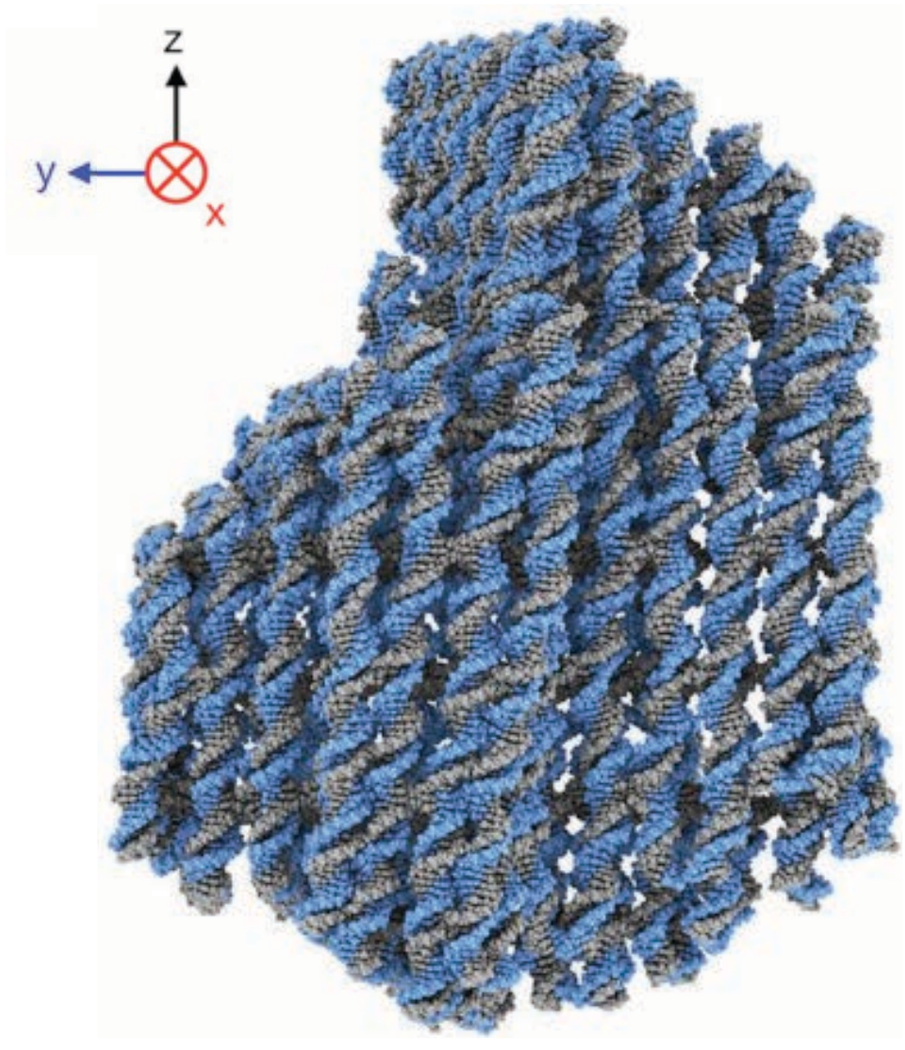


Bai *et al*, PNAS 109:20012 (2012)

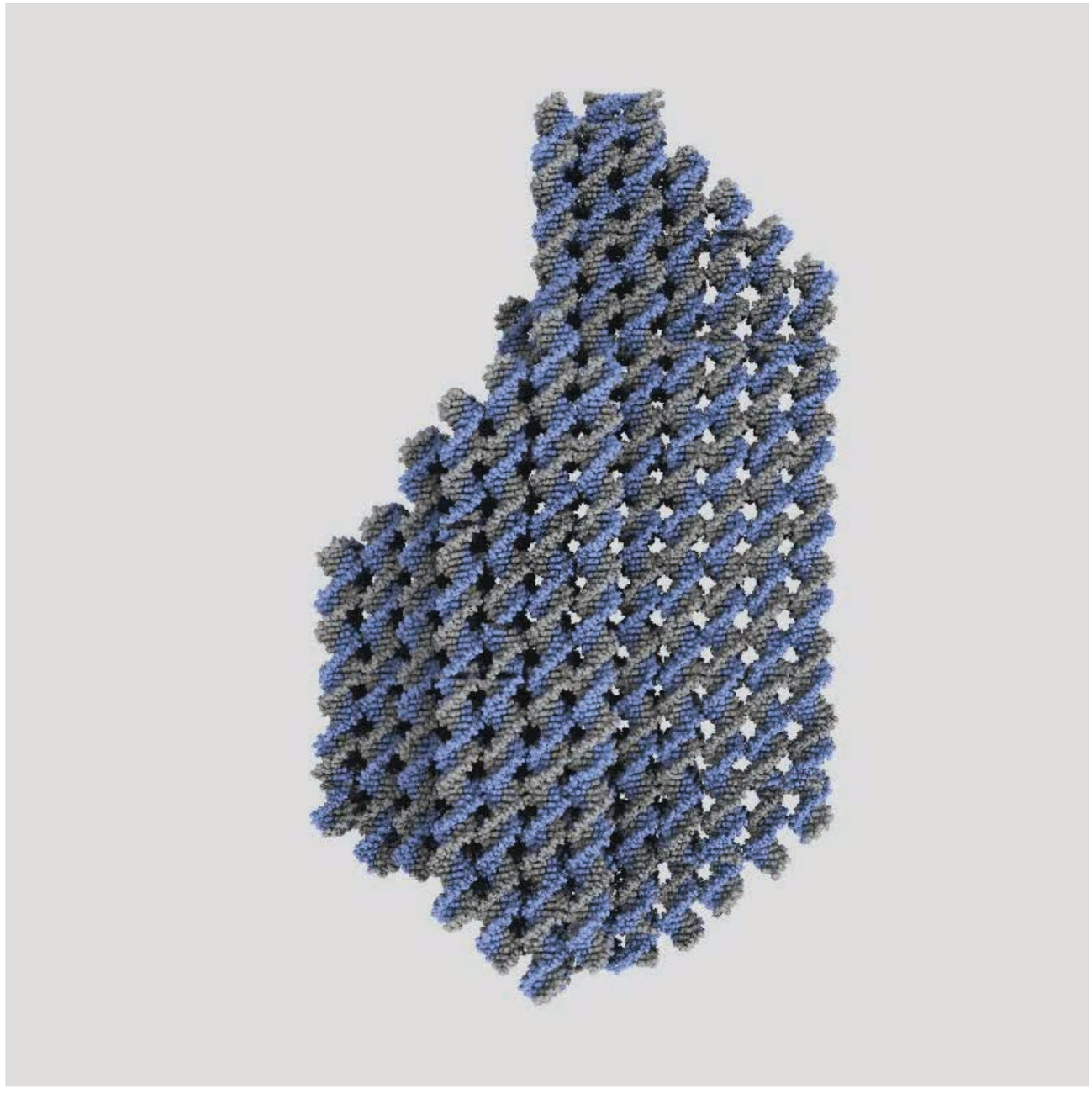


7M atom solvated model
~200 ns MD trajectory

MD simulation of the cryo-EM object starting from a caDNAno design

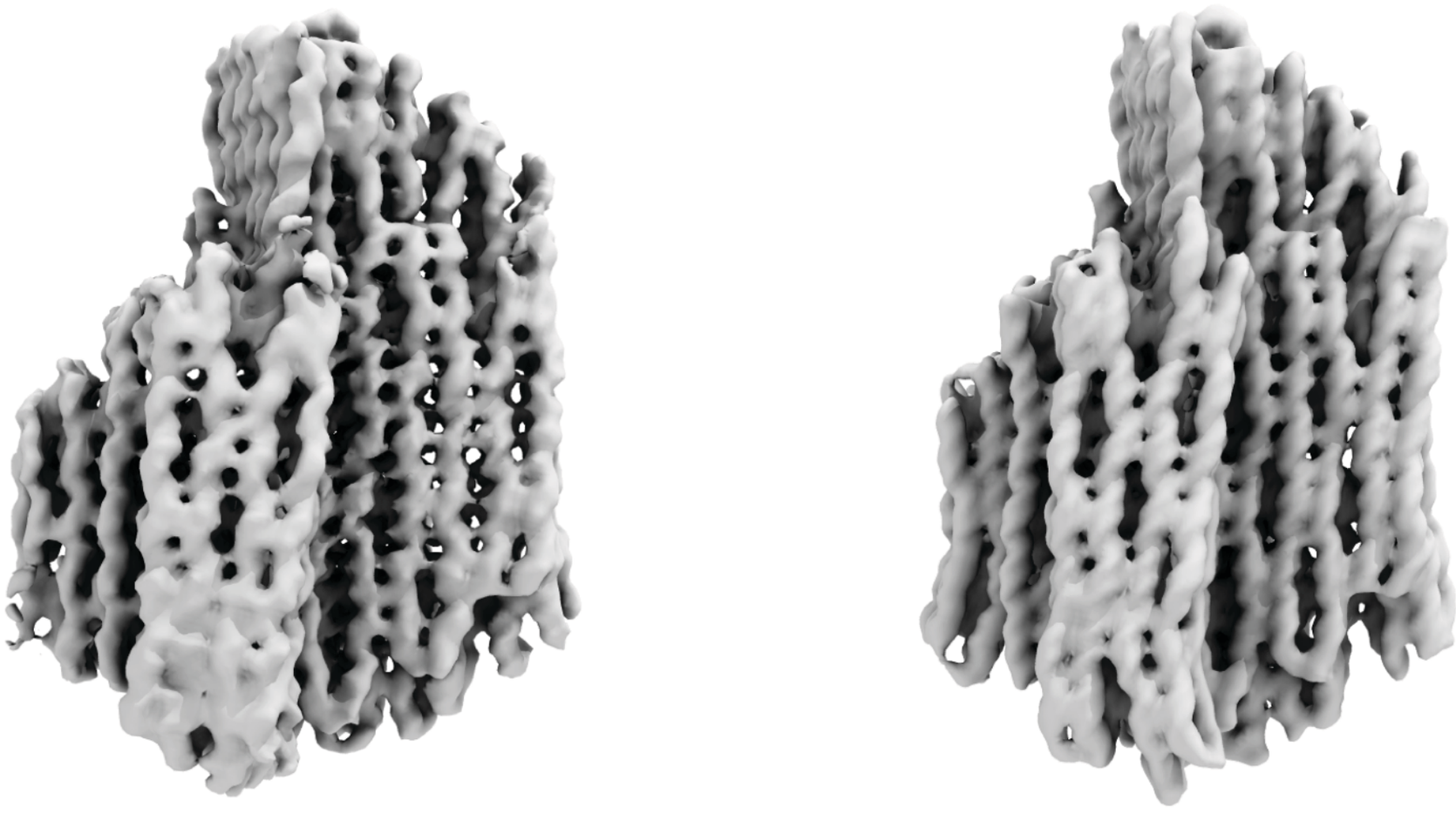


Bai *et al*, PNAS 109:20012 (2012)

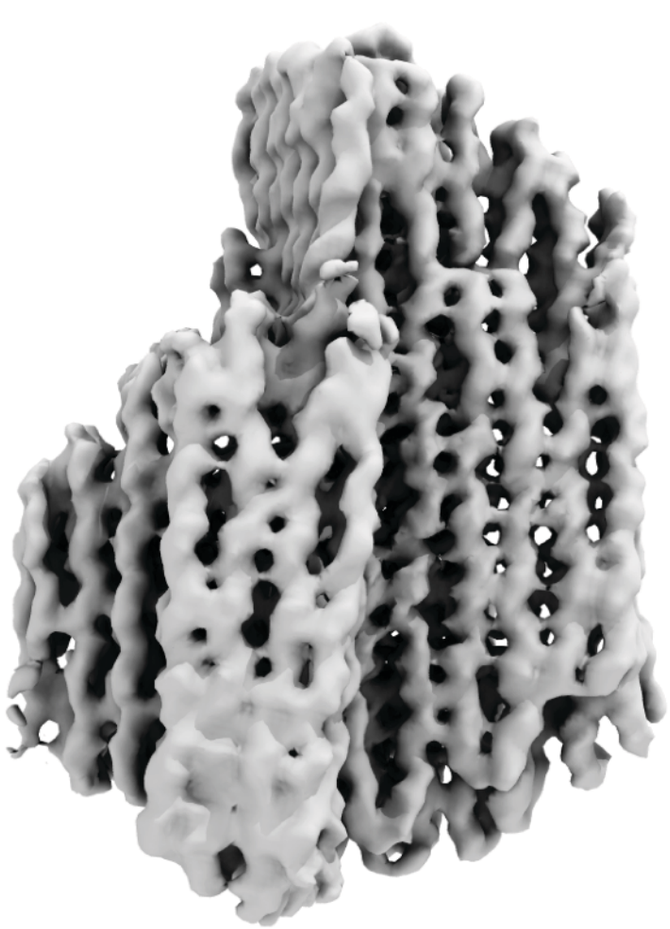


7M atom solvated model
~200 ns MD trajectory

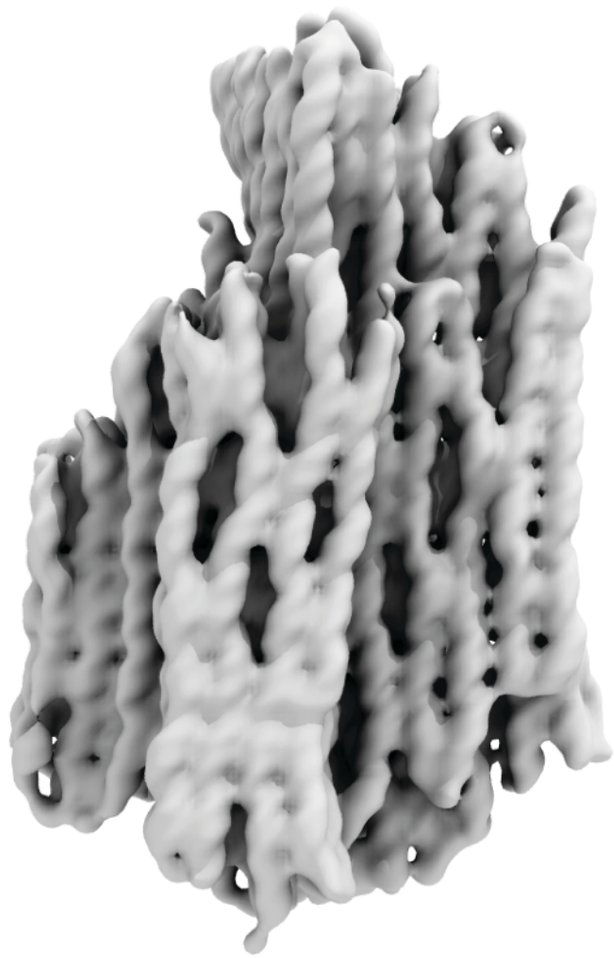
Simulated electron density is similar to experimental electron density



Electron density maps



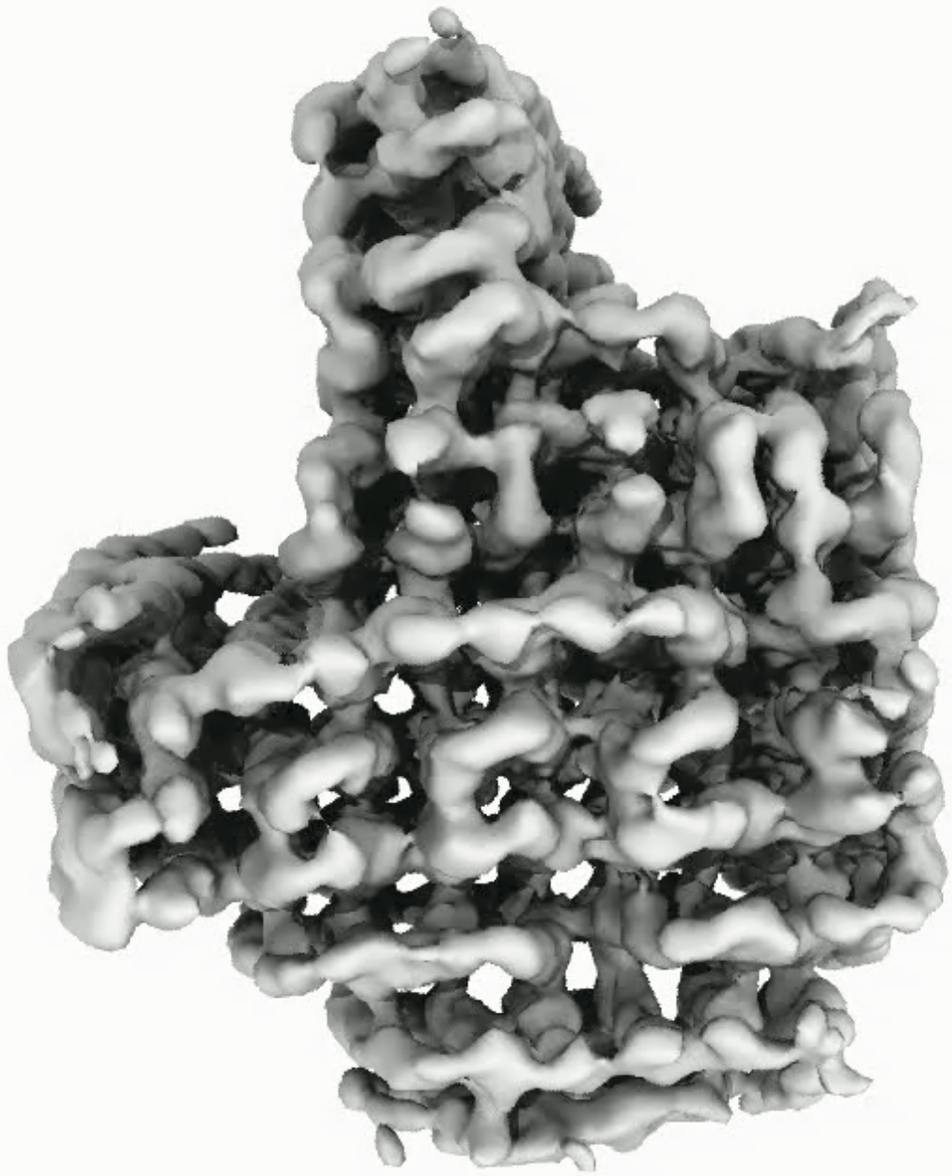
Cryo-EM reconstruction



All-atom MD simulation

Comparison with experiment

Maffeo, Yoo & Aksimentiev, *NAR 44: 3013 (2016)*



EM density psuedo-atomic model simulation

ENRG MD for Origami Structur ×

bionano.physics.illinois.edu/origami-structure

The Aksimentiev Group

Theoretical and Computational Research at the Interface of Physics, Biology, and Nanotechnology

ENRG MD For Origami Structure Prediction

Upload a DNA origami design .json file

Choose File No file chosen

Select the origami lattice. *

- (•) Square
- ( ) Honeycomb

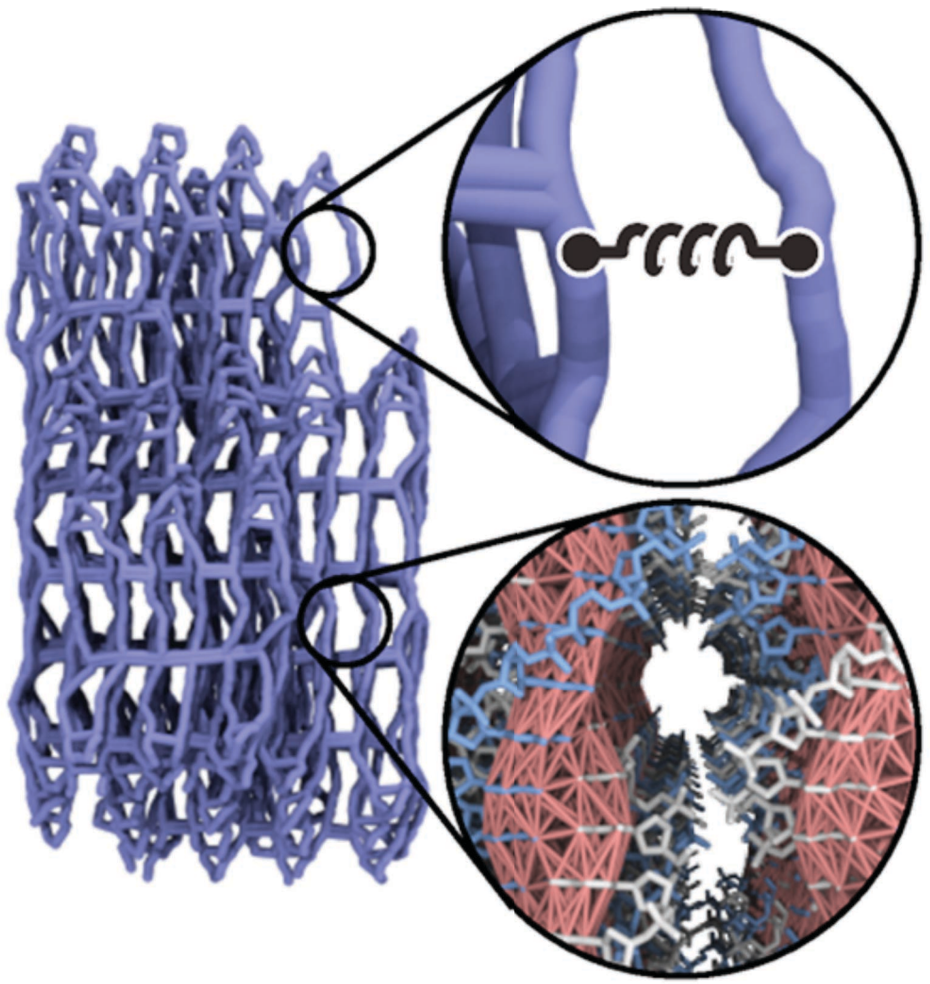
Select the scaffold sequence. *

- (•) m13mp18 (up to 7,249 bases)
- ( ) Custom

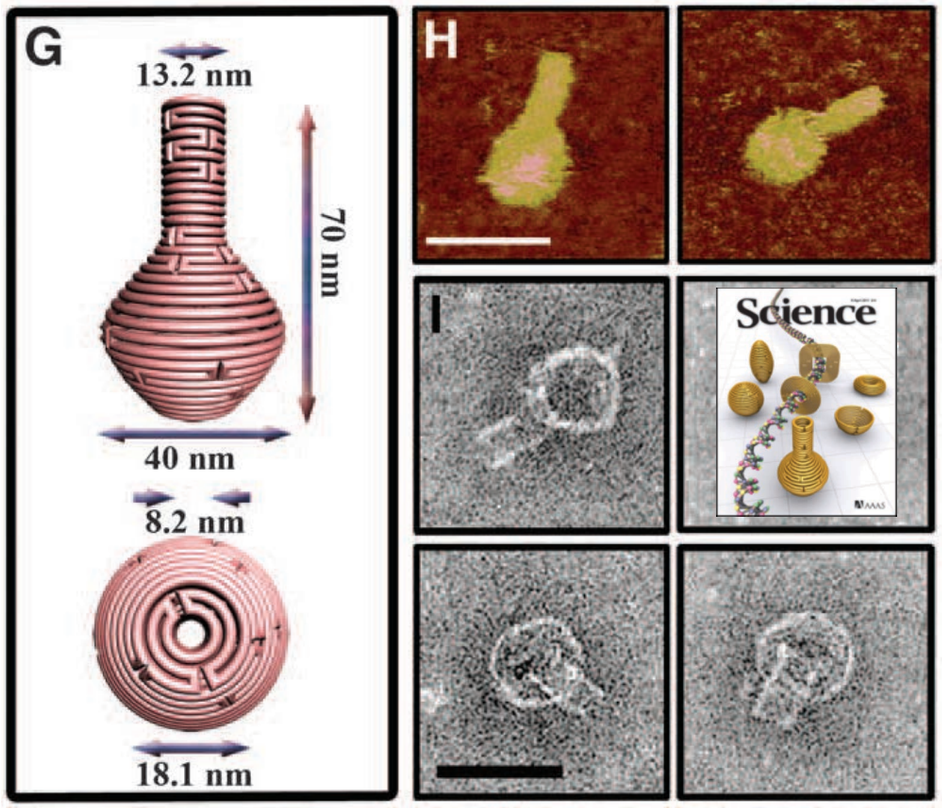
Simulation package *

- (•) NAMD (CHARMM FF)
- ( ) Gromacs (AMBER FF; beta coming soon)

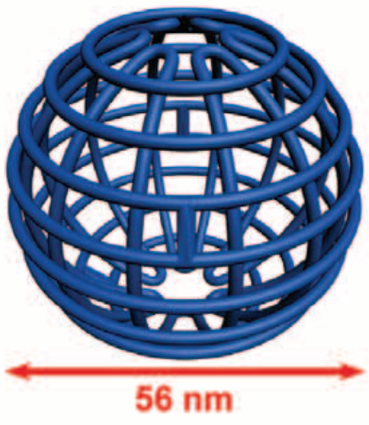
Create simulation files



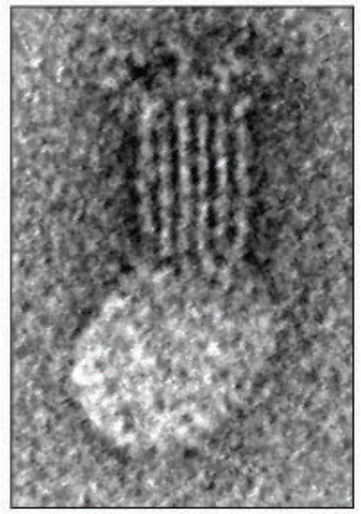
DNA origami structures

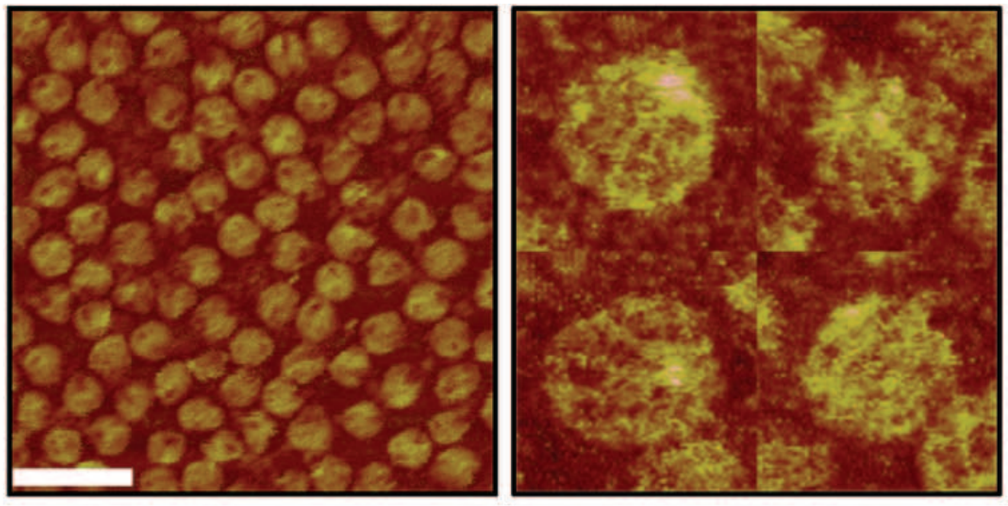


Yan and coworkers, Science (2011)

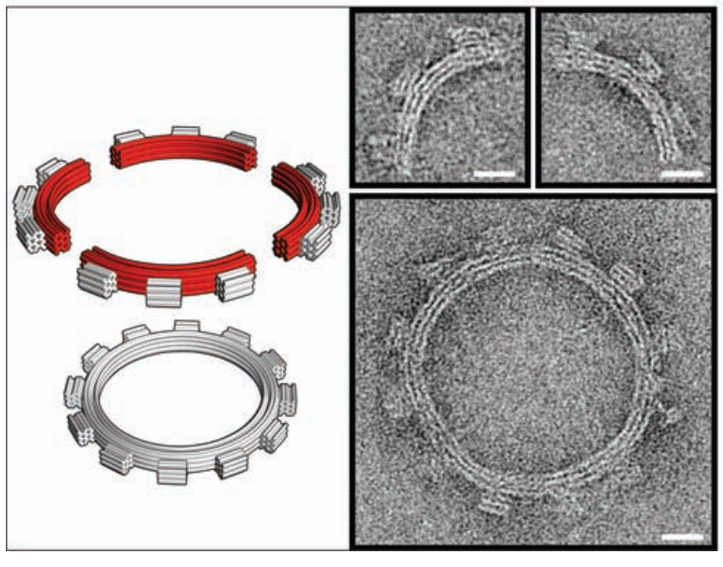


Dietz and coworkers, Science (2012)

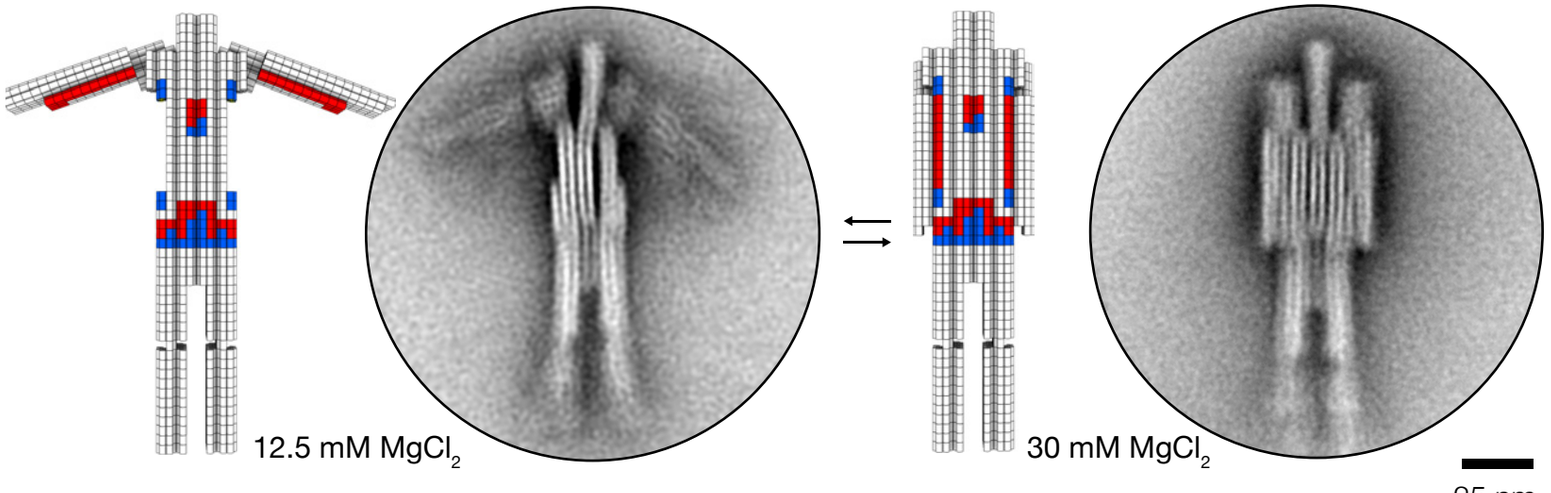




Yan and coworkers, Science (2013)



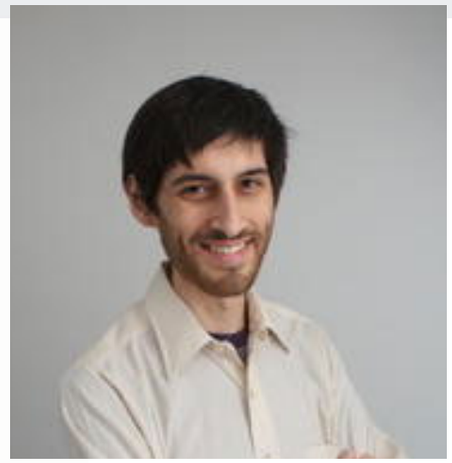
Shih and coworkers, Science (2009)



Dietz and coworkers, Science (2015)

Multi-Resolution DNA (mrDNA) model

Nucleic Acids Research 48: 5135 (2020)

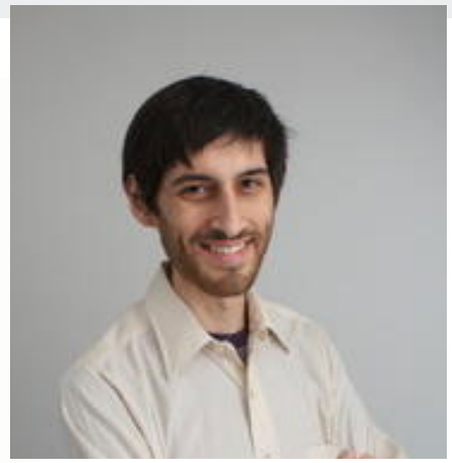


Chris Maffeo

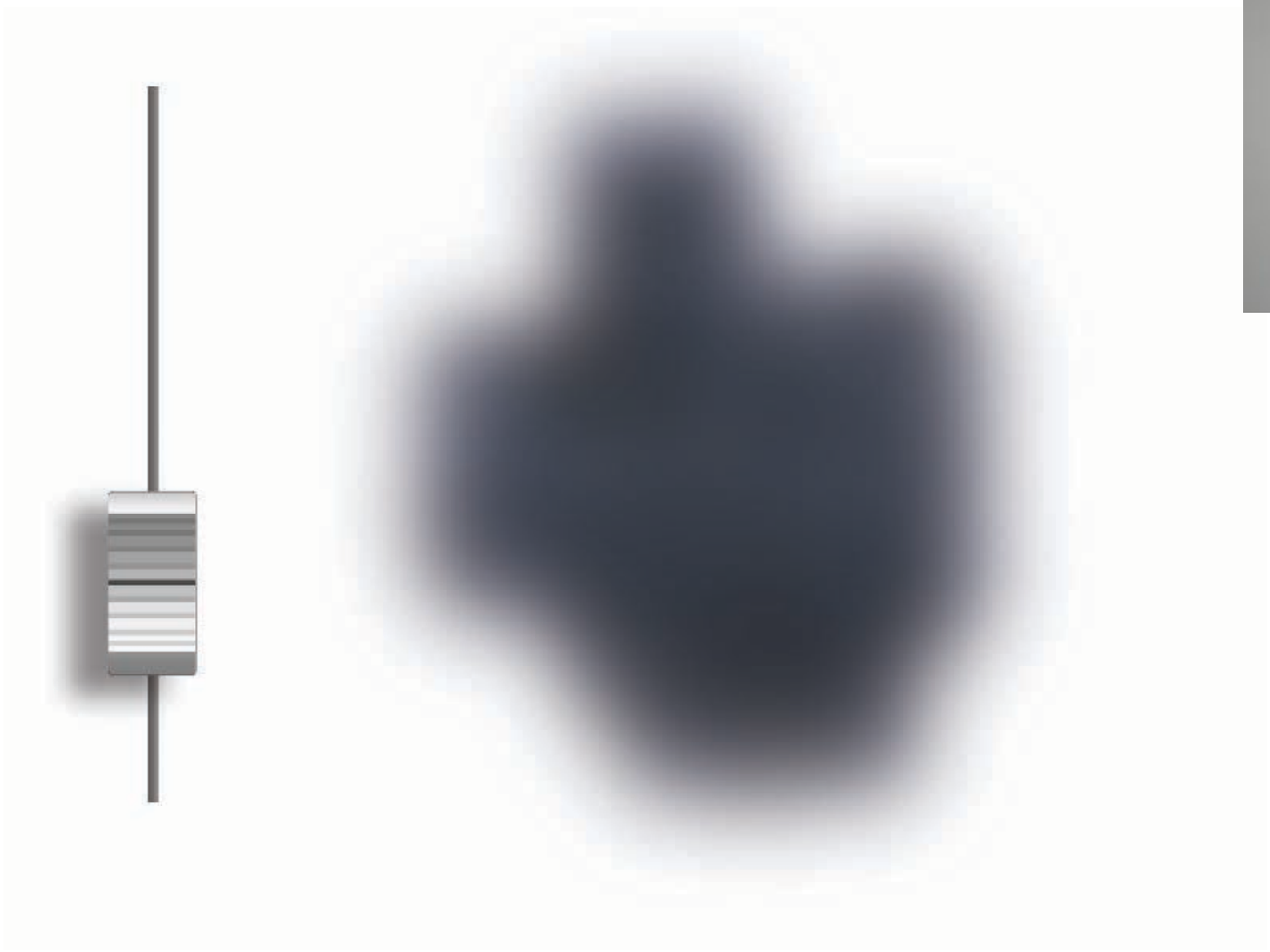
Strategy: change resolution for speed and detail

Multi-Resolution DNA (mrDNA) model

Nucleic Acids Research 48: 5135 (2020)



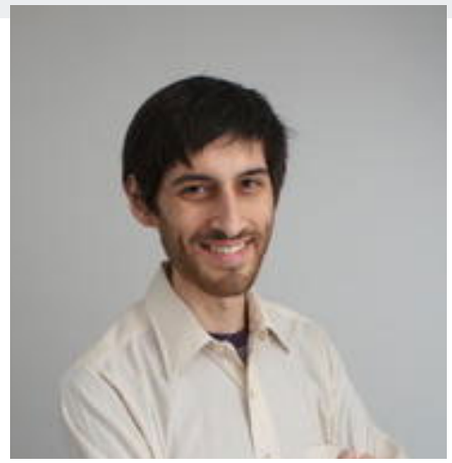
Chris Maffeo



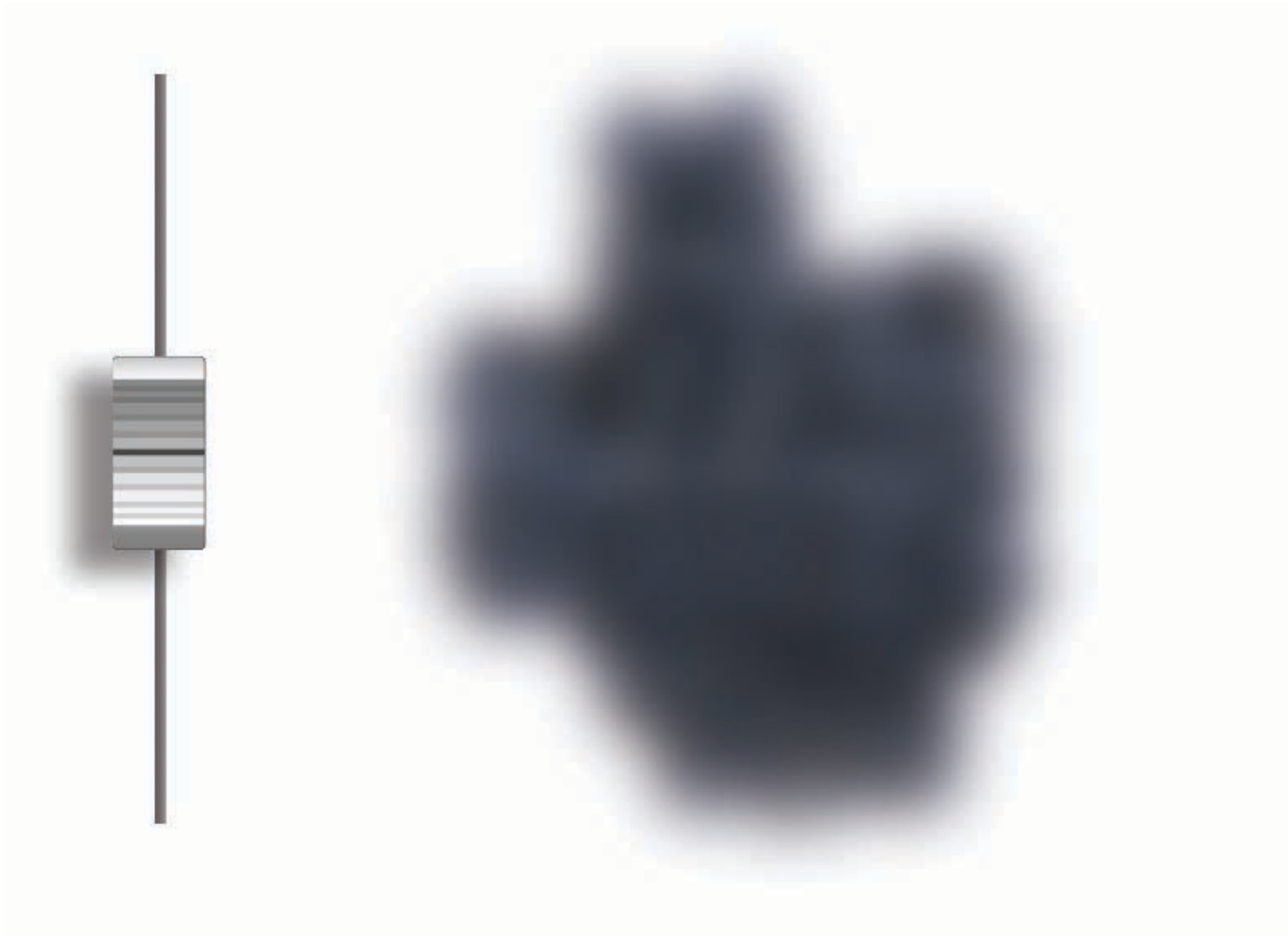
Strategy: change resolution for speed and detail

Multi-Resolution DNA (mrDNA) model

Nucleic Acids Research 48: 5135 (2020)



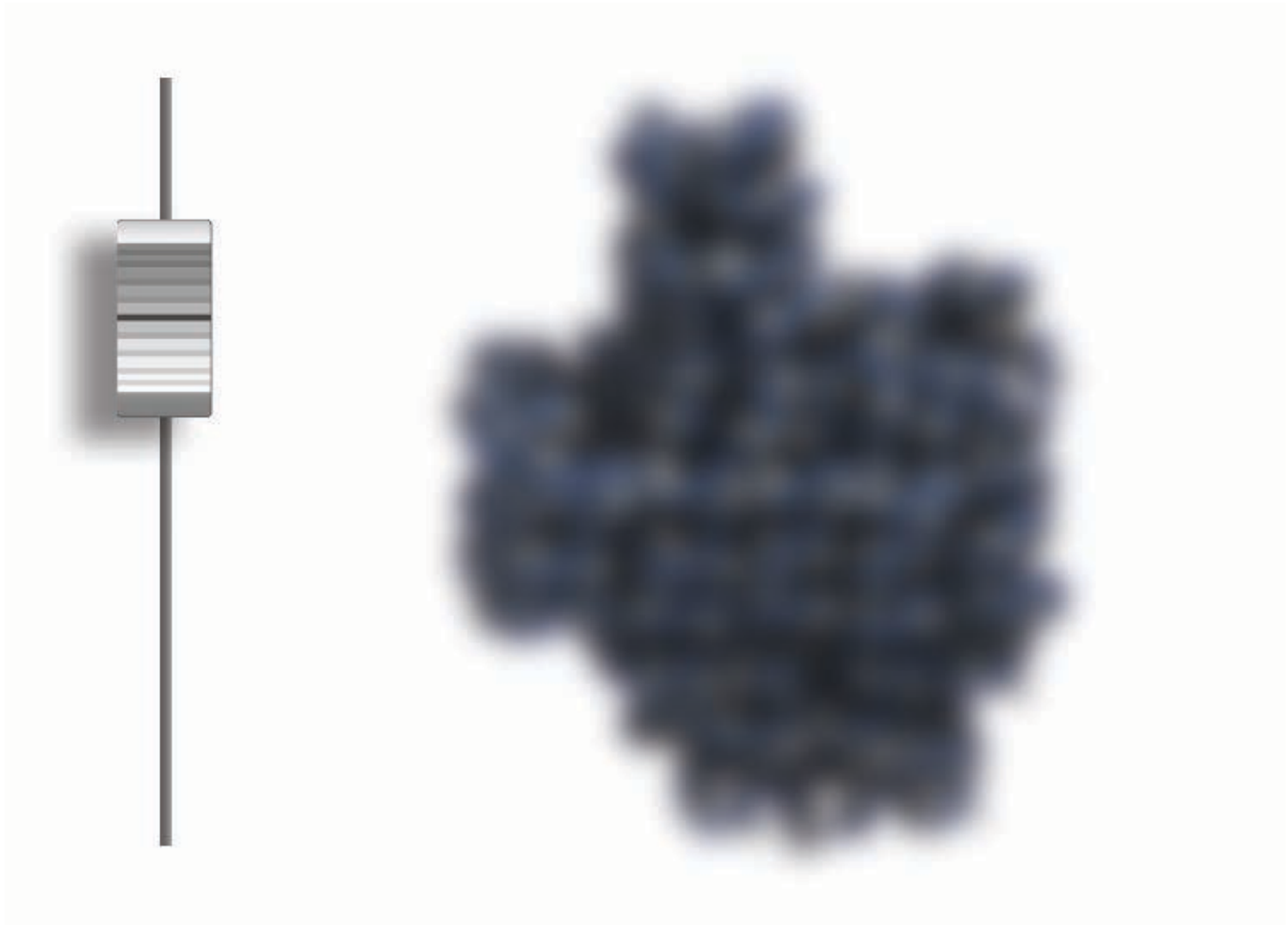
Chris Maffeo

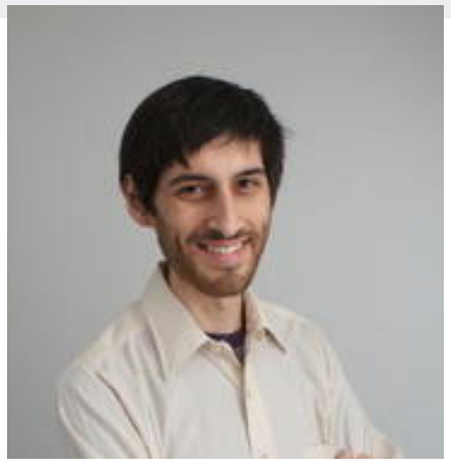


Strategy: change resolution for speed and detail

Multi-Resolution DNA (mrDNA) model

Nucleic Acids Research 48: 5135 (2020)



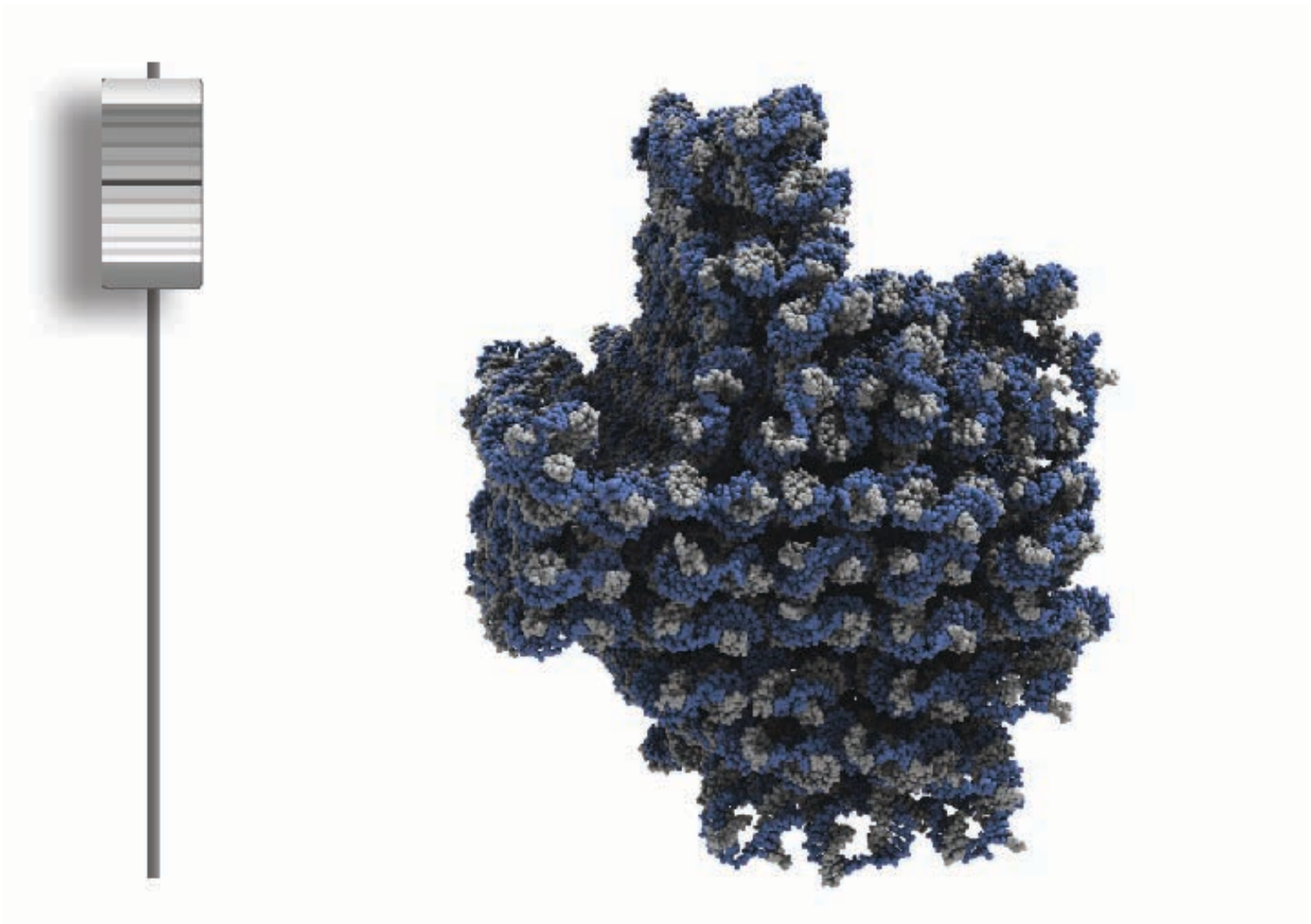


Chris Maffeo

Strategy: change resolution for speed and detail

Multi-Resolution DNA (mrDNA) model

Nucleic Acids Research 48: 5135 (2020)

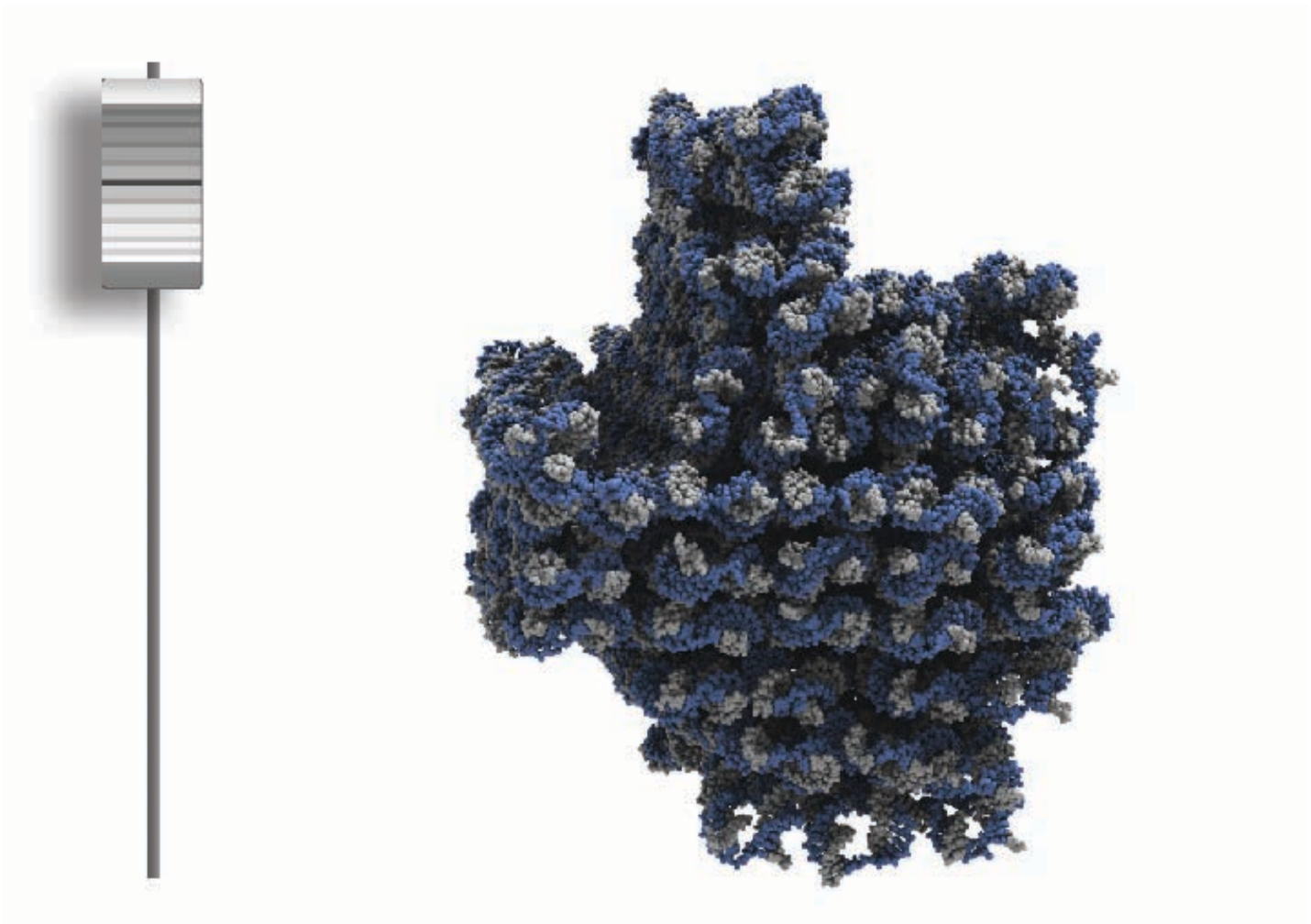


Chris Maffeo

Strategy: change resolution for speed and detail

Multi-Resolution DNA (mrDNA) model

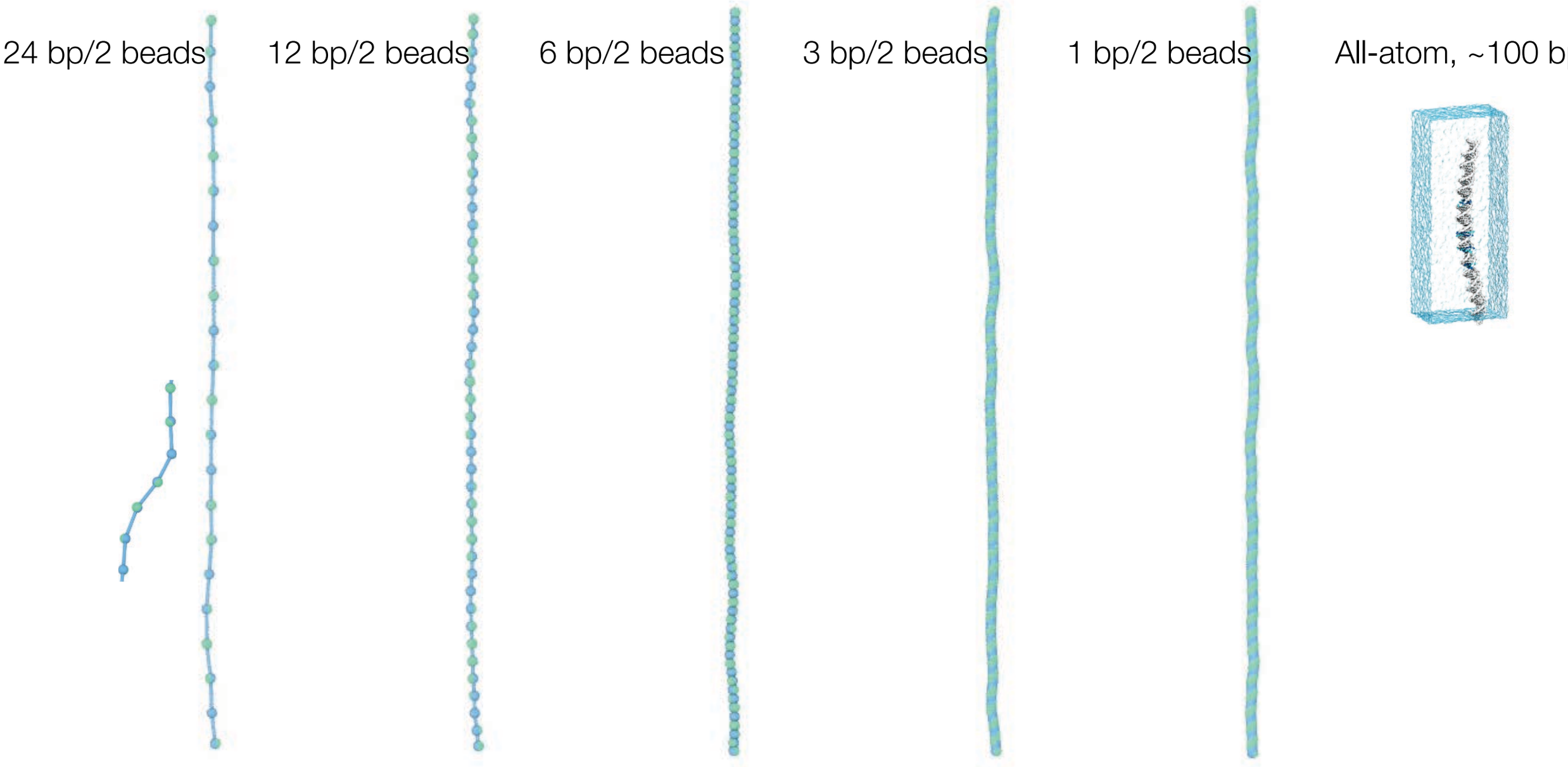
Nucleic Acids Research 48: 5135 (2020)



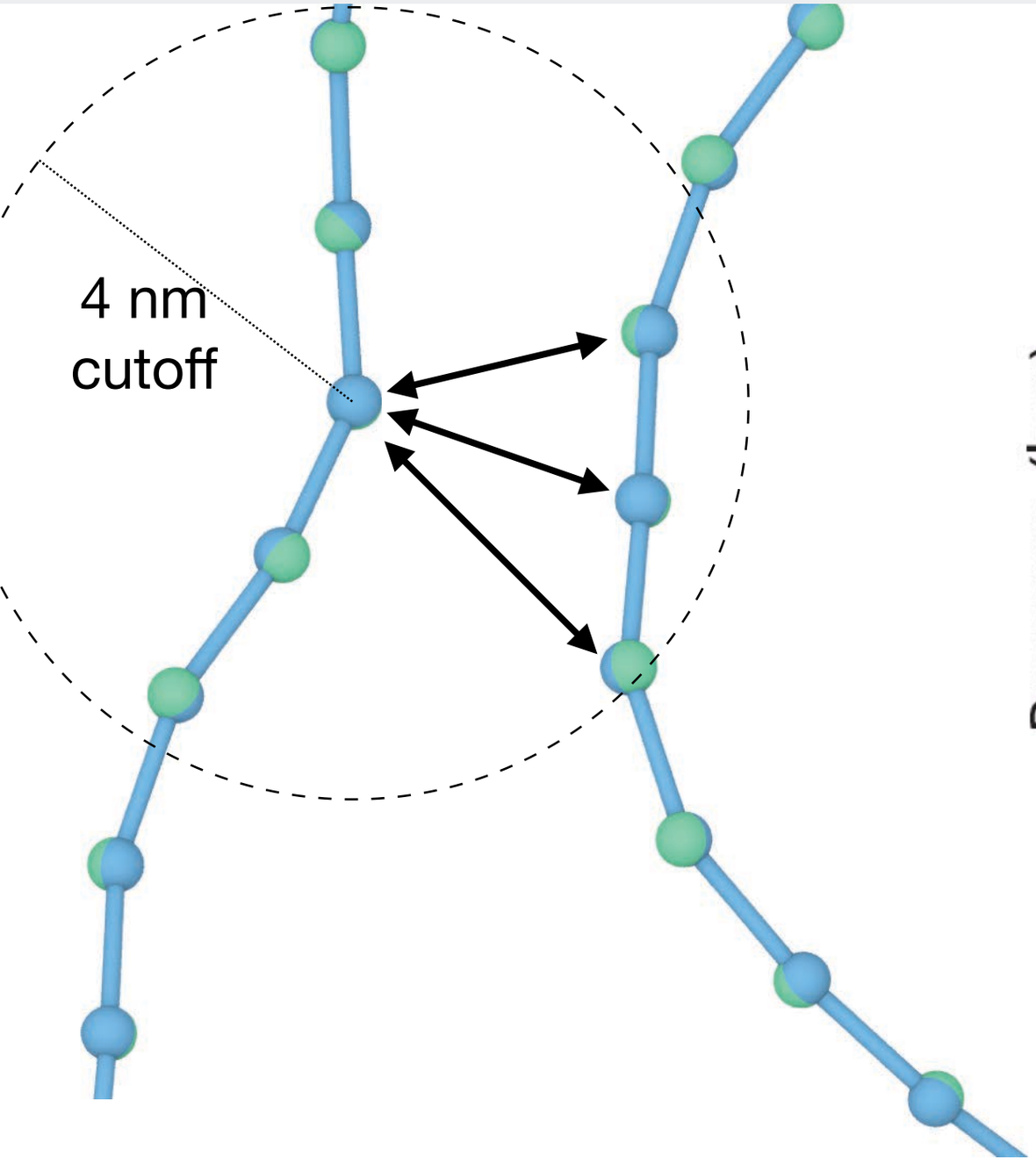
Chris Maffeo

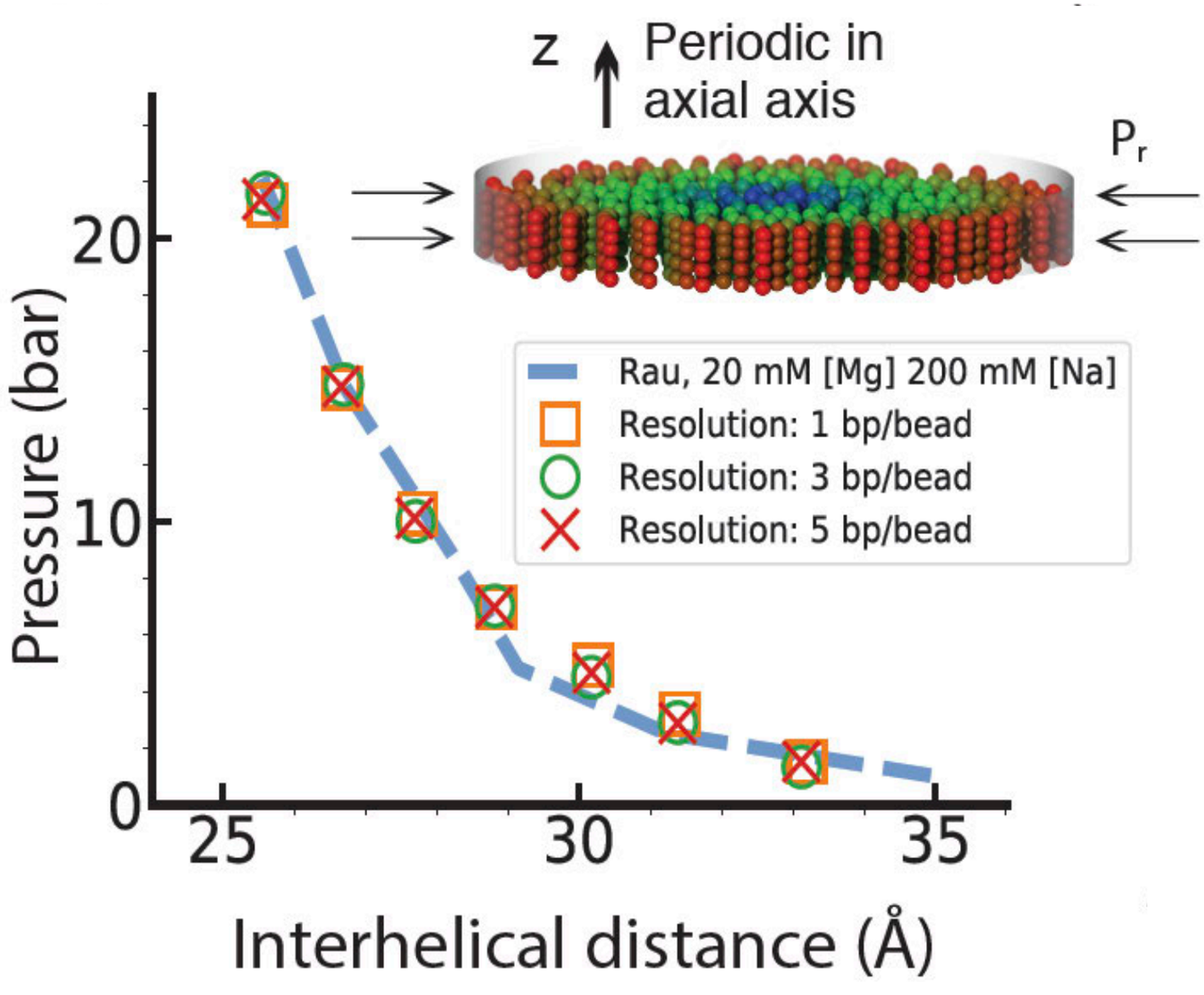
Strategy: change resolution for speed and detail

500 bp dsDNA fragment modeled at different resolutions

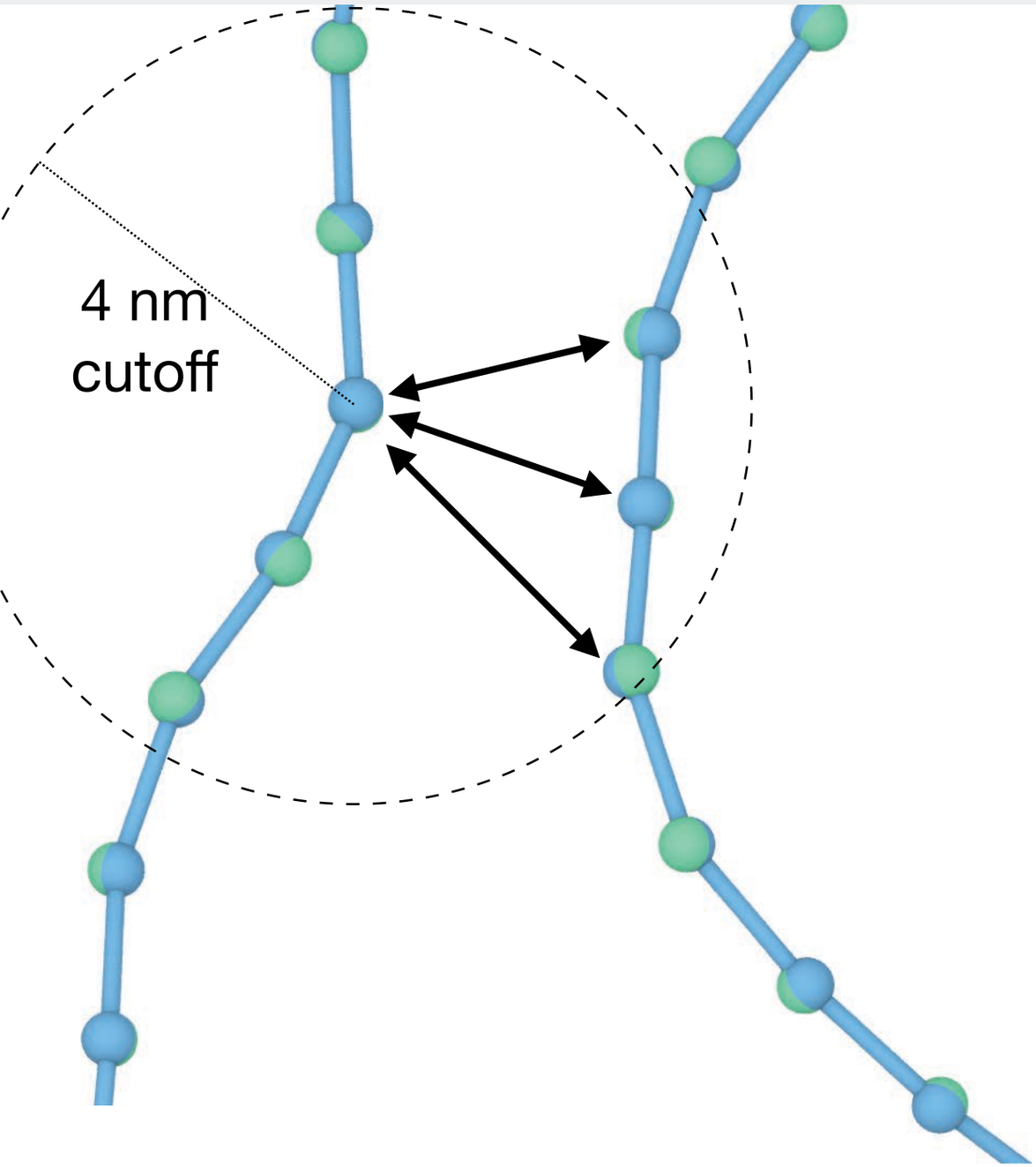


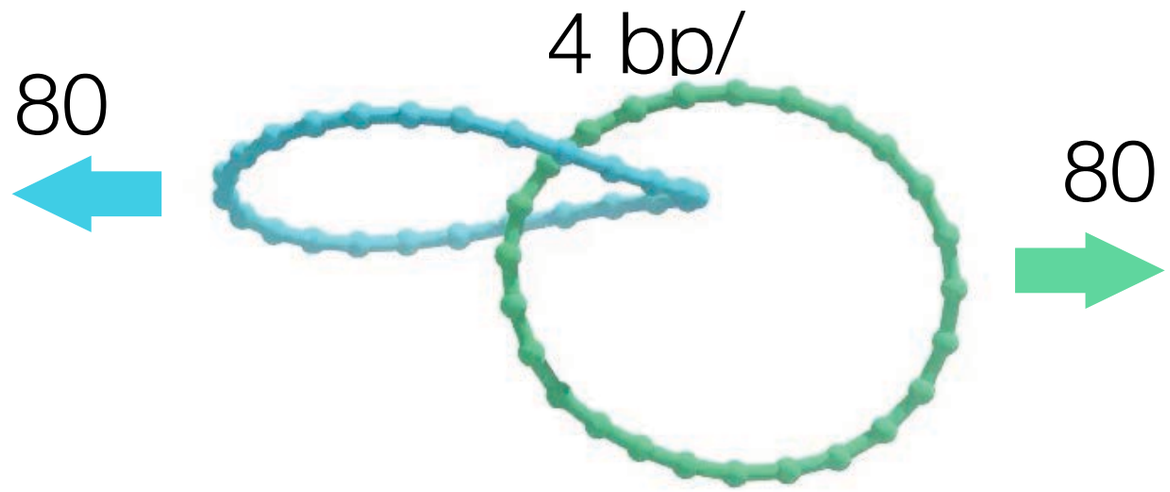
Interactions in a simple coarse-grained DNA model

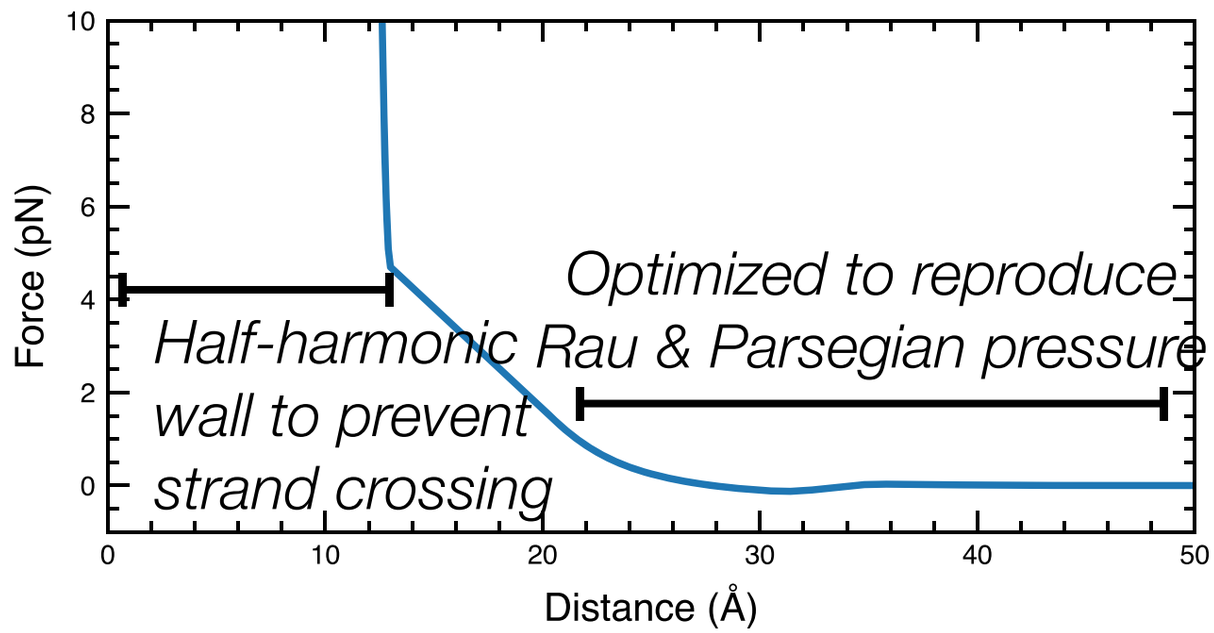




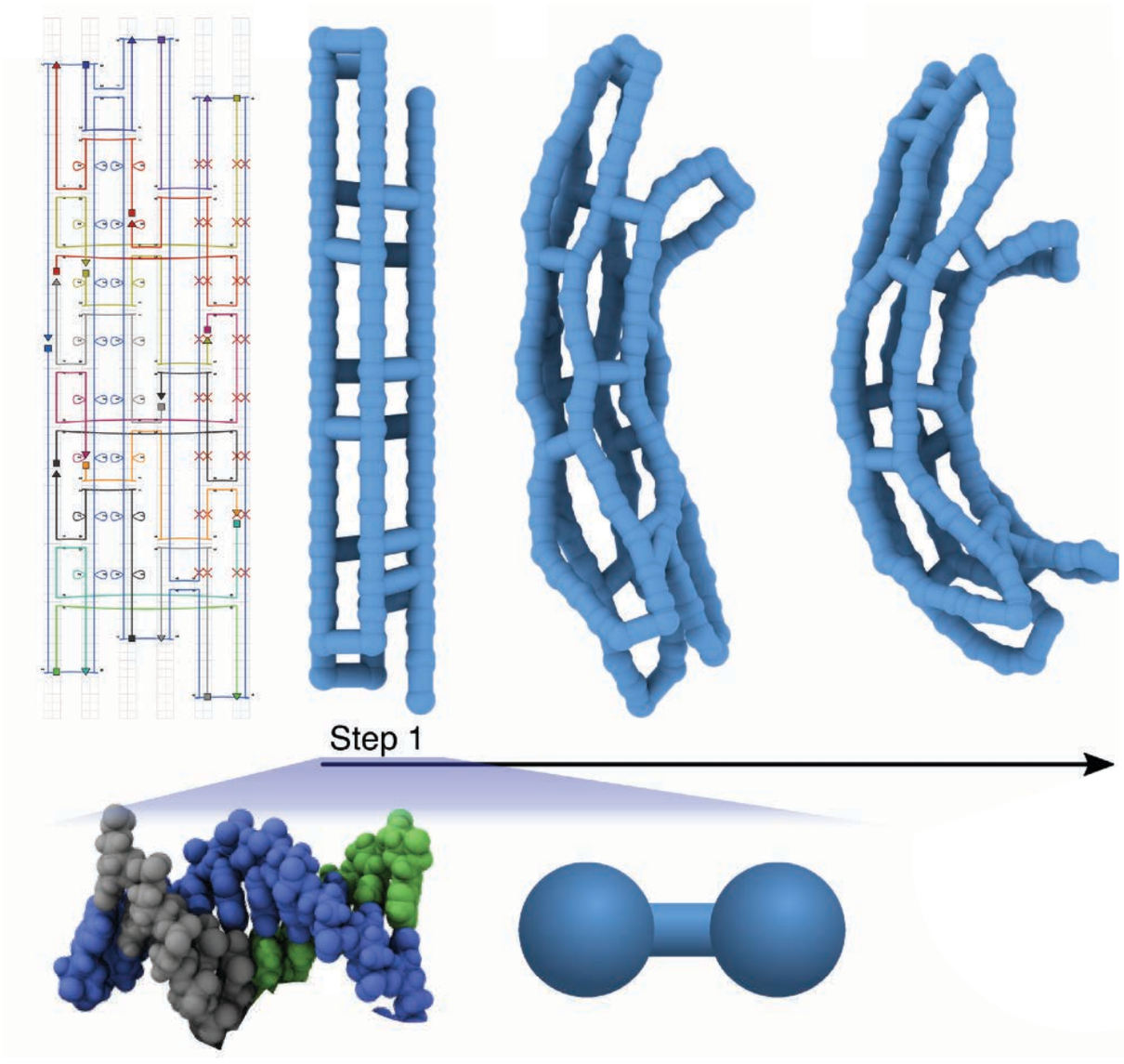
Interactions in a simple coarse-grained DNA model





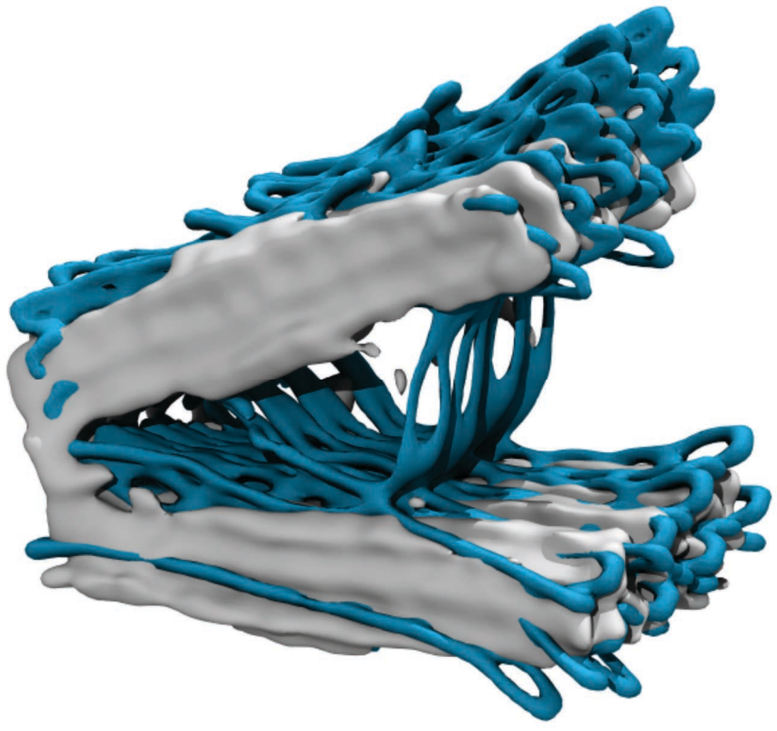


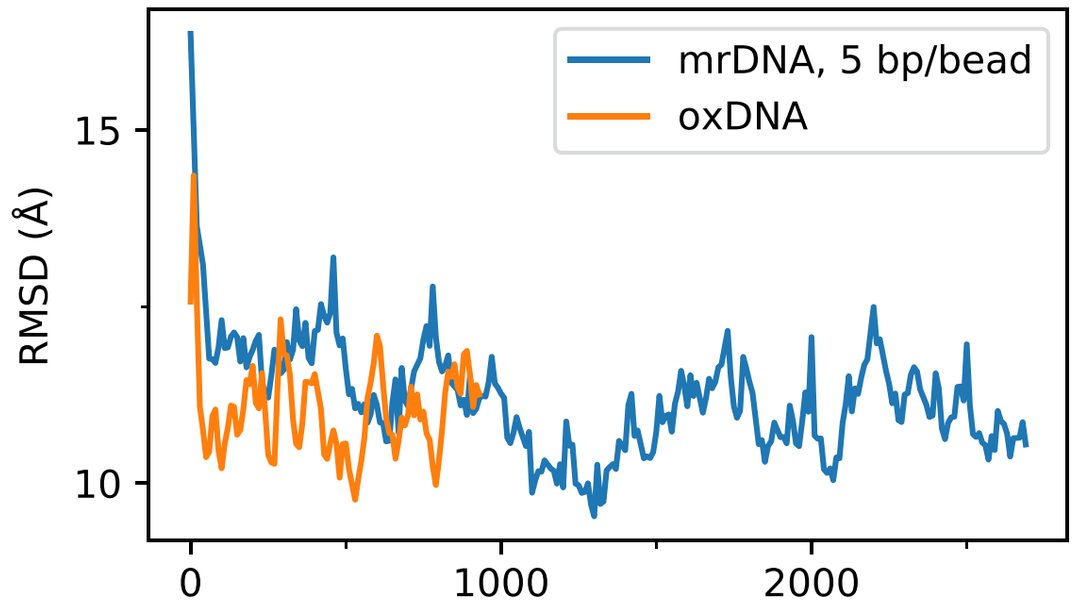
Typical structural relaxation procedure



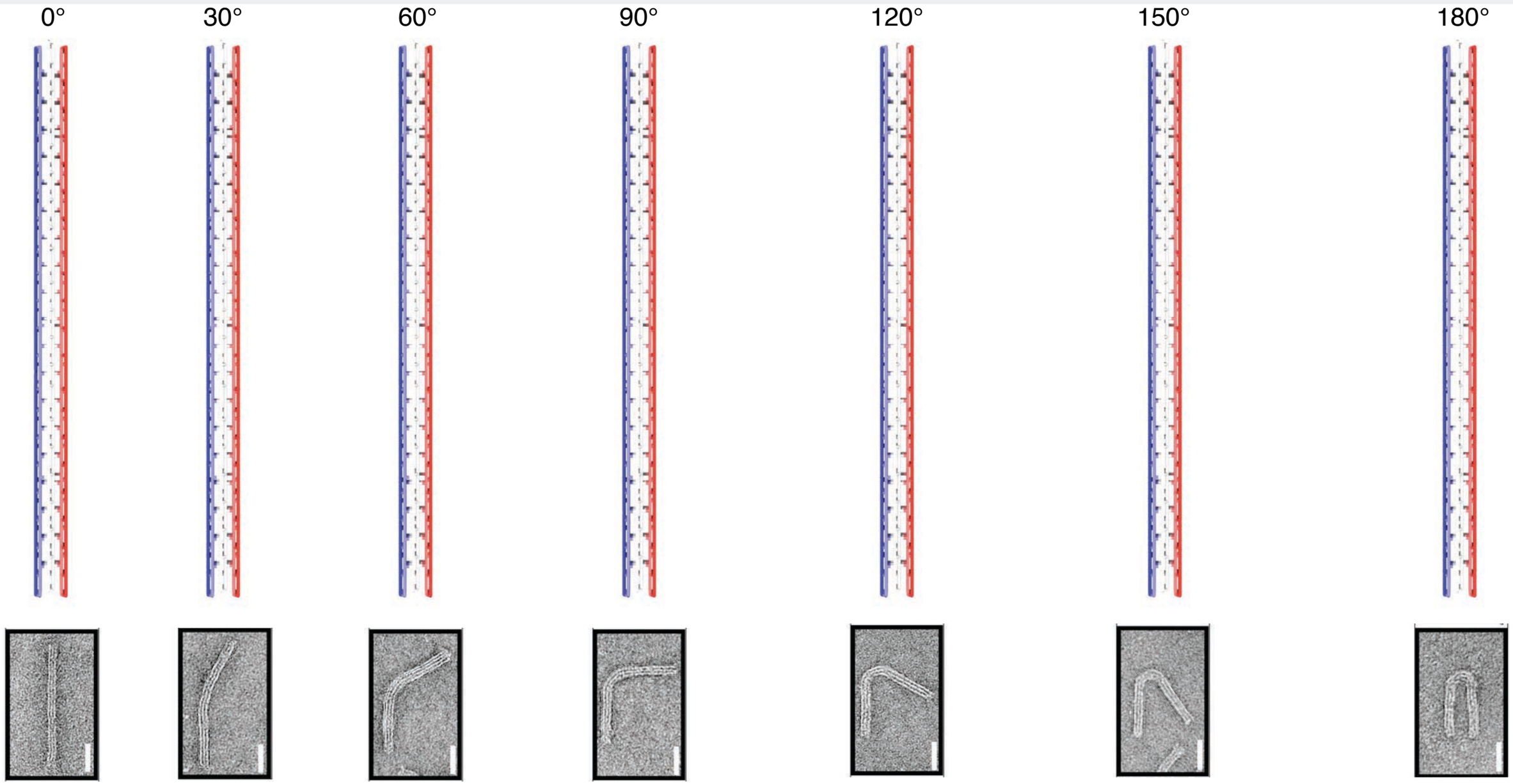
Nucleic Acids Research 48: 5135 (2020)

Multi-resolution simulations provide highly detailed structures quickly





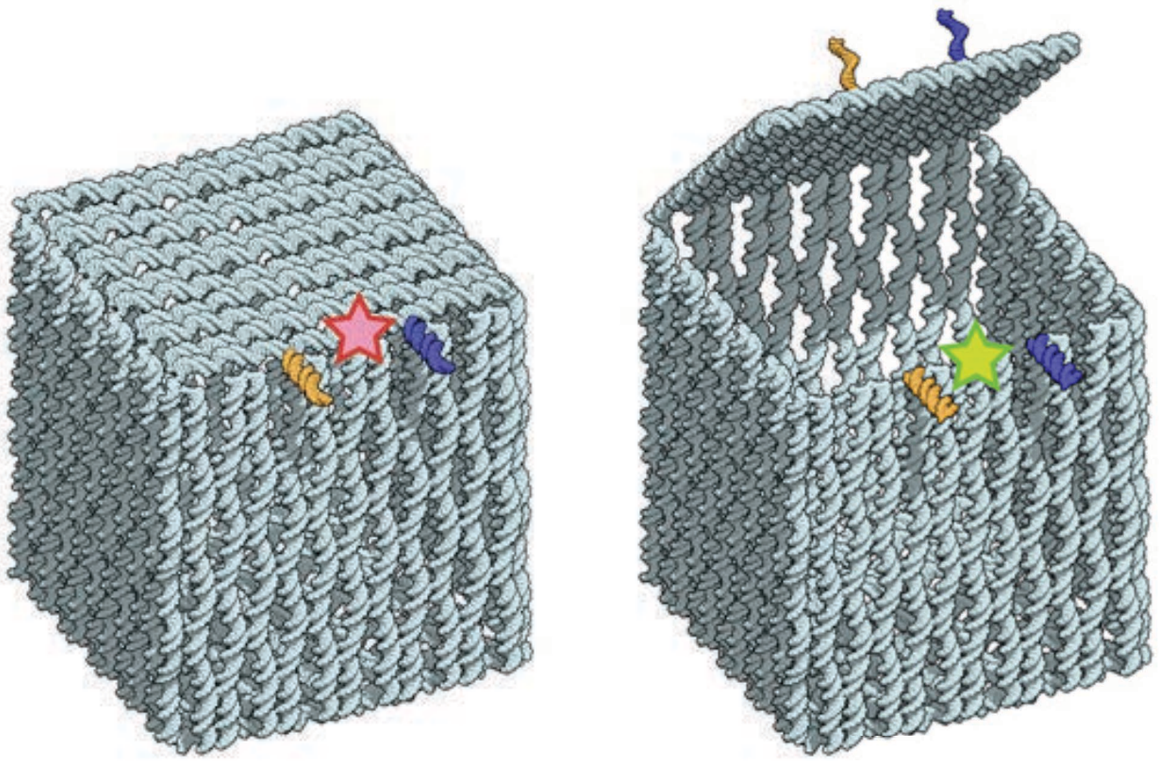
Coarse-grained model captures programmed curvature



Experiment from : Science 325:725

Adaptive resolution simulation of DNA origami systems

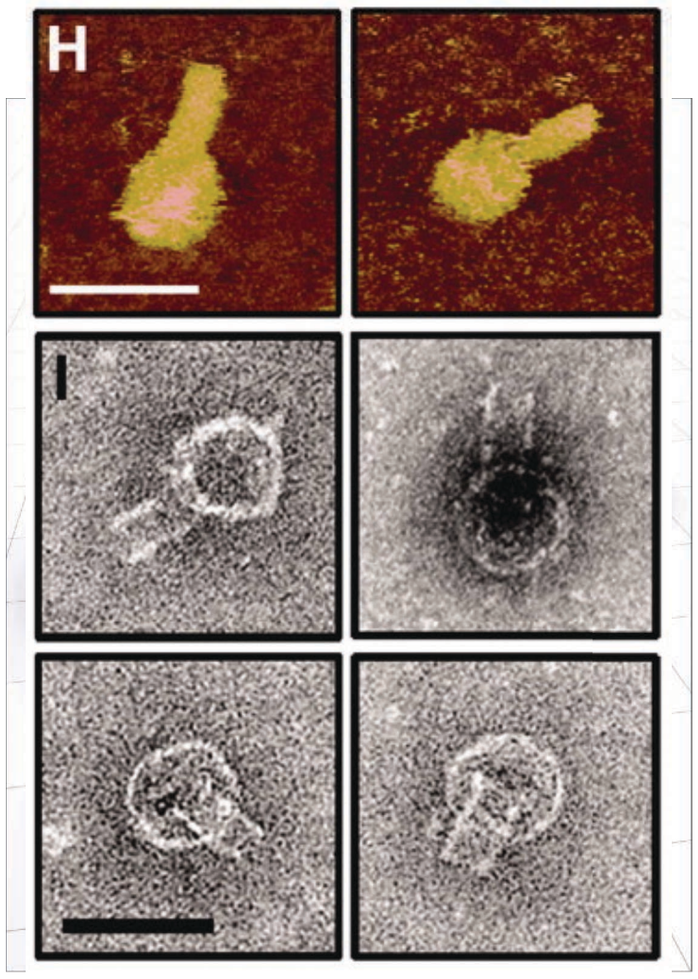
Victoria Birkedal Group



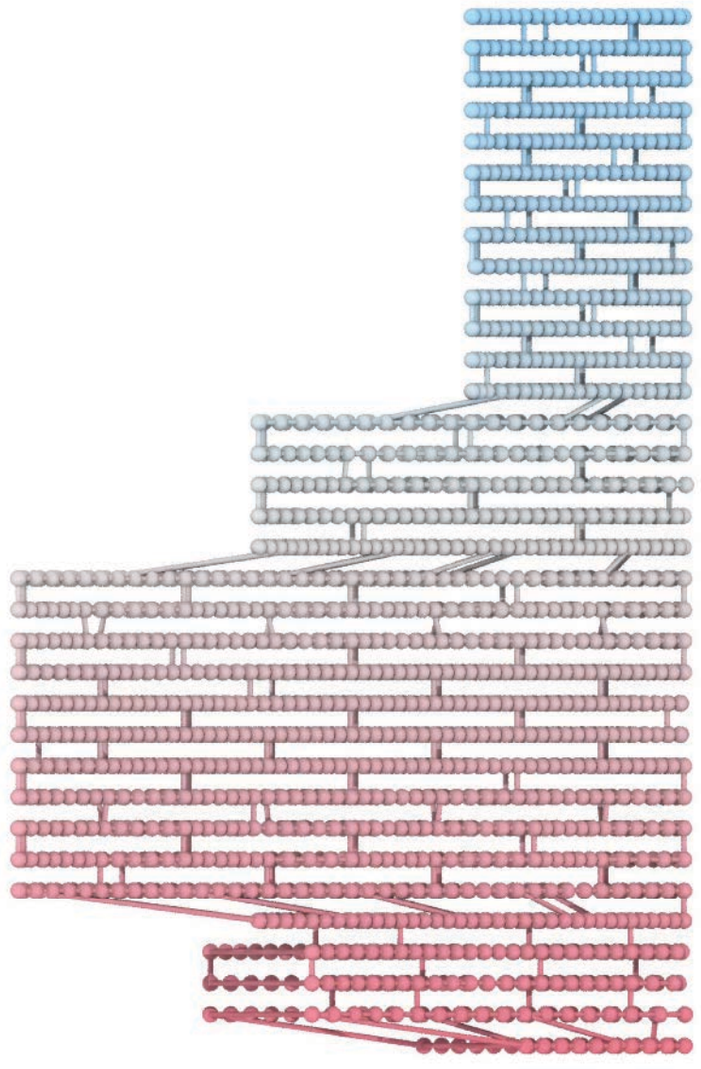
Andersen et al., Nature 2009

Used to interpret FRET characterization of DNA box variants: *Nanoscale* 11:18475 (2019)

Multi-resolution modeling of self-assembled DNA nanostructures

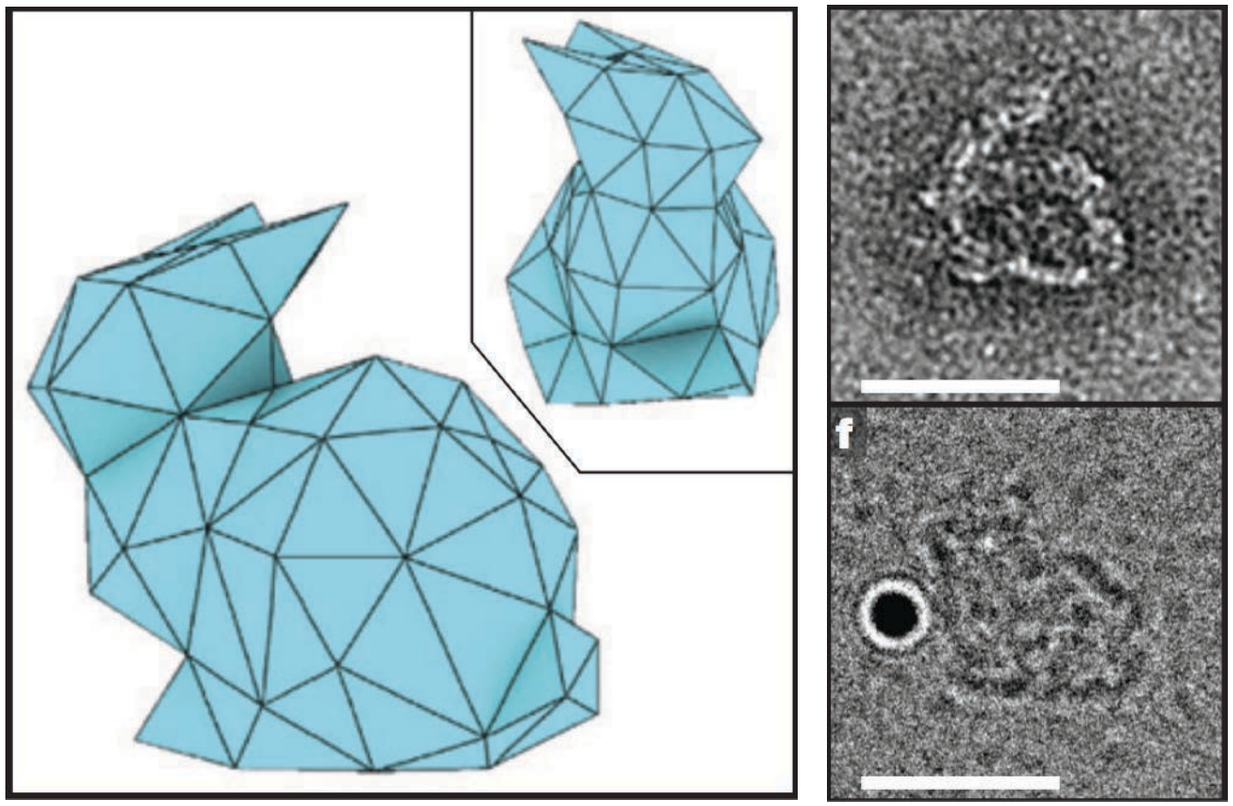


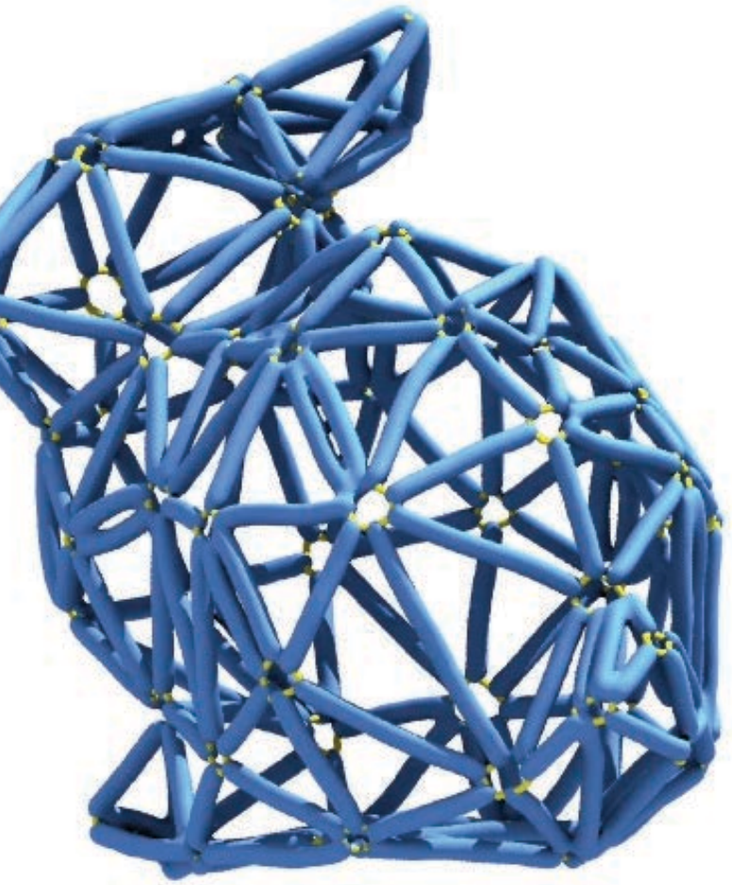
Dongran Han, Suchetan Pal, Jeanette Nangreave, Zhengtao Deng, Yan Liu, Hao Yan
Science 332:342



Nucleic Acids Research 48: 5135 (2020)

Multi-resolution workflow extended to DNA polyhedral meshes

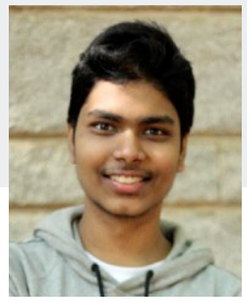




E Benson, A Mohammed, J Gardell, S Masich, E Czeizler, P Orponen & B Högberg
Nature 523:441

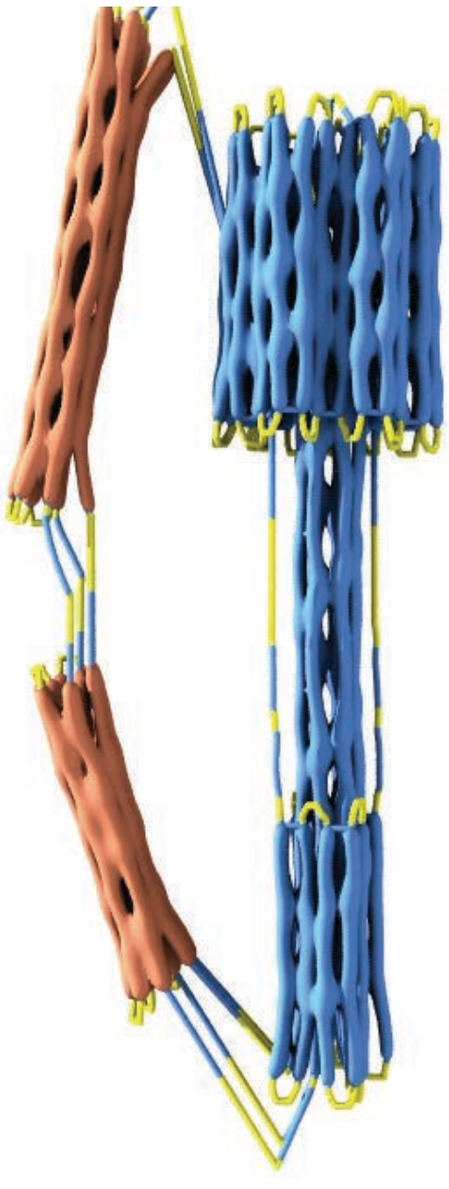
Nucleic Acids Research 48: 5135 (2020)

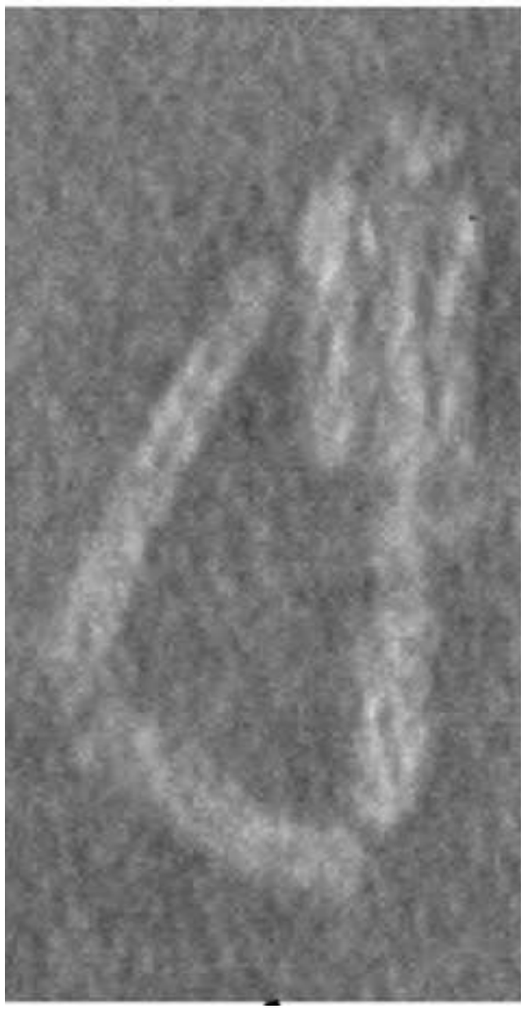
Coarse-grained simulations for sampling structural fluctuations

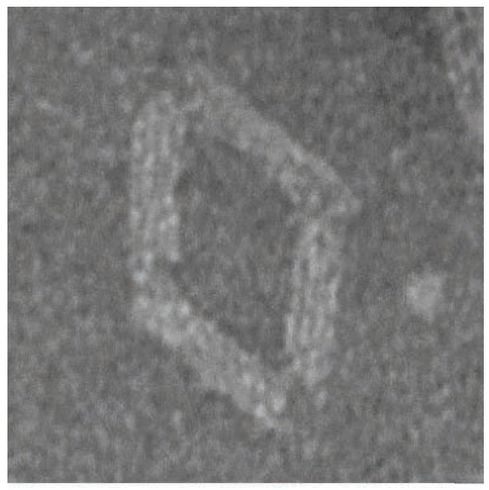


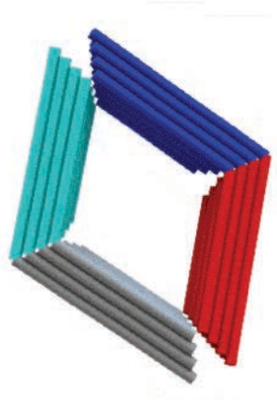
Kumar Sarthak

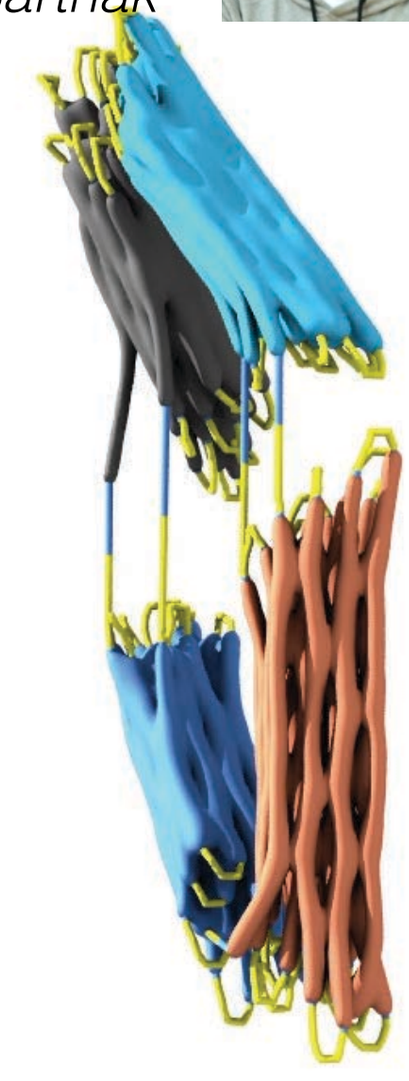
Alexander E. Marras, Lifeng Zhou, Hai-Jun Su and Carlos E. Castro PNAS 112:713





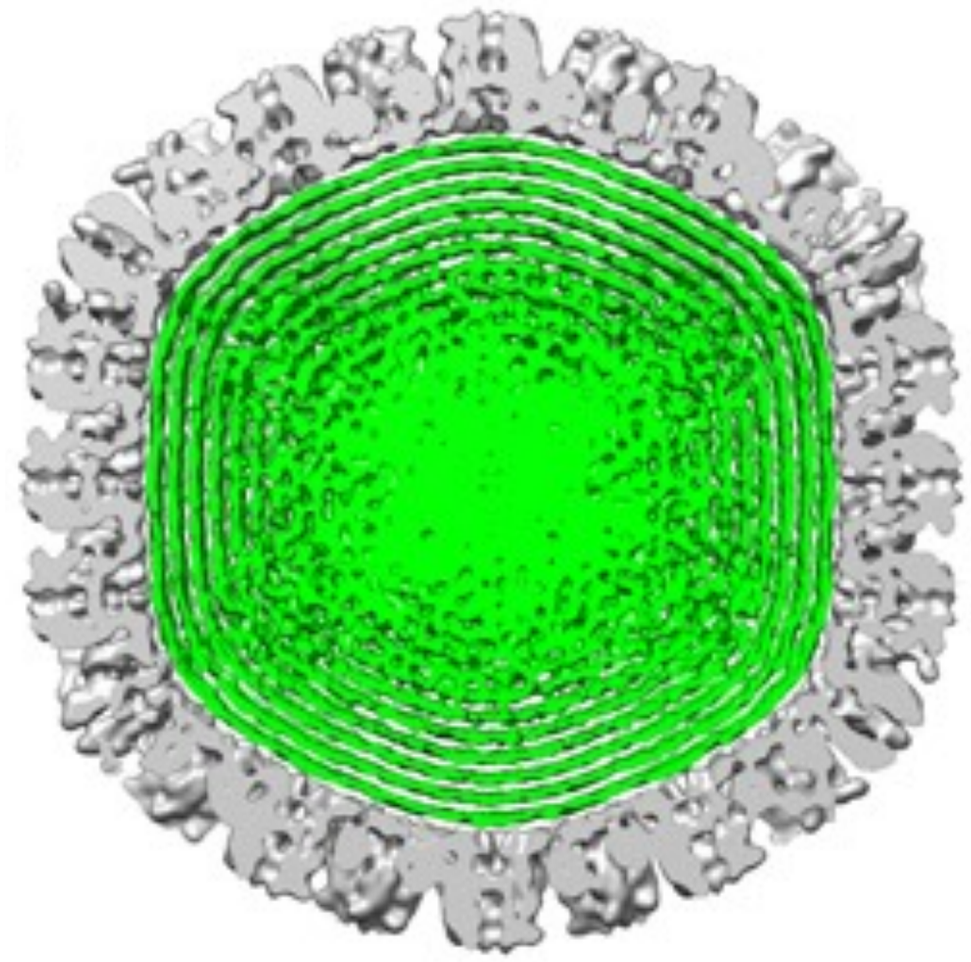






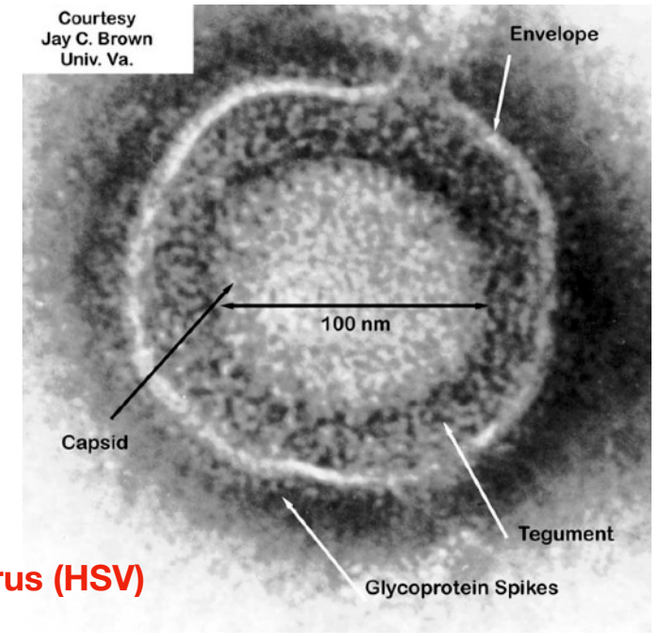
Nucleic Acids Research 48: 5135 (2020)

Viral genome, the program of infection

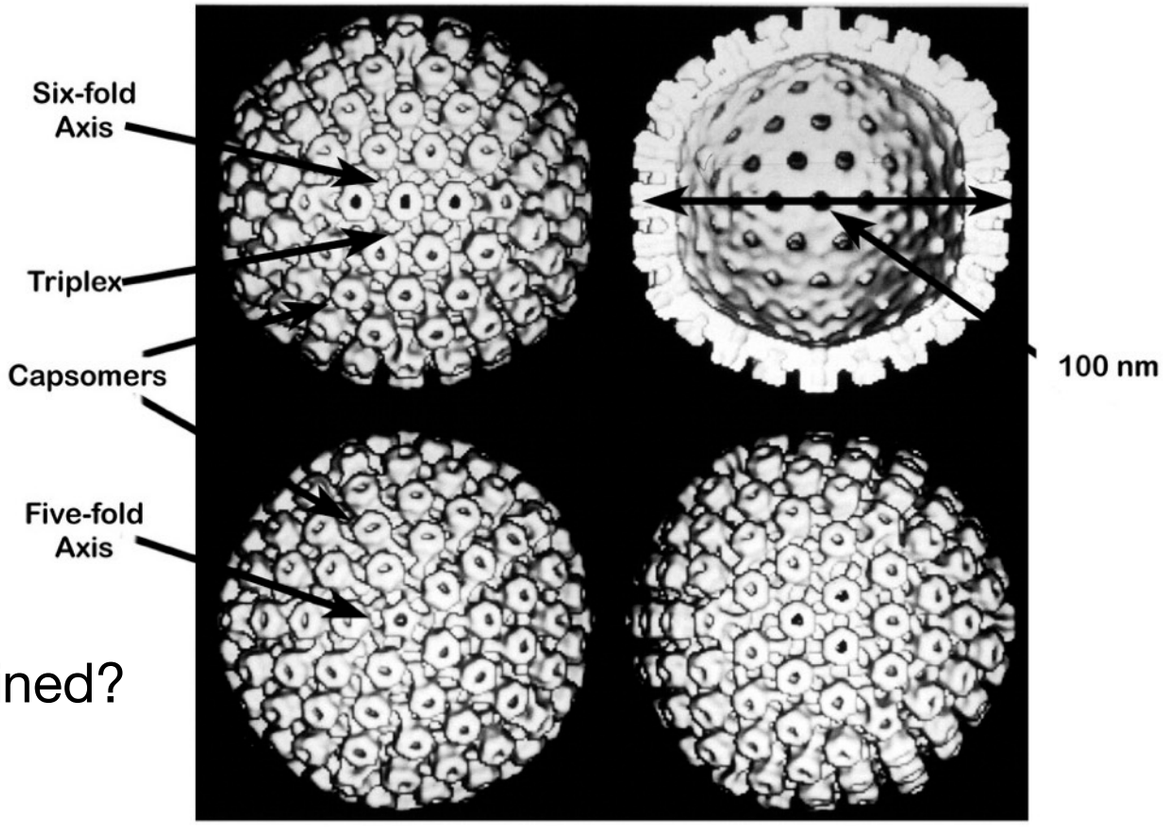


Cryoem reconstruction with concentric rings
(Evilevitch et al, UIUC)

DNA is a highly charged polymer!



Herpes virus (HSV)



http://darwin.bio.uci.edu/~faculty/wagner/hsv2f.html

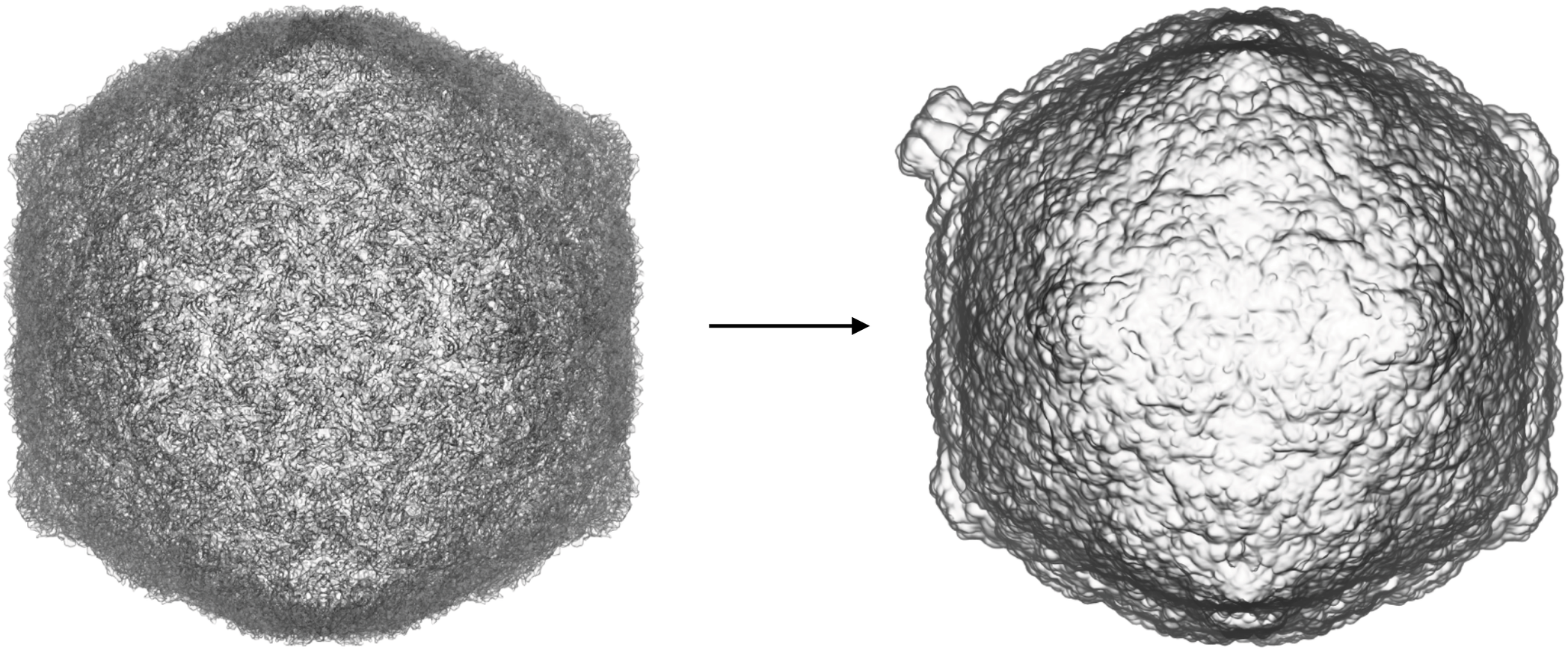
Open questions:

- What is the 3D structure of the genome?
- How genome ejection is triggered and sustained?
- Can it be used as a drug target?

Packaging viruses with ARBD

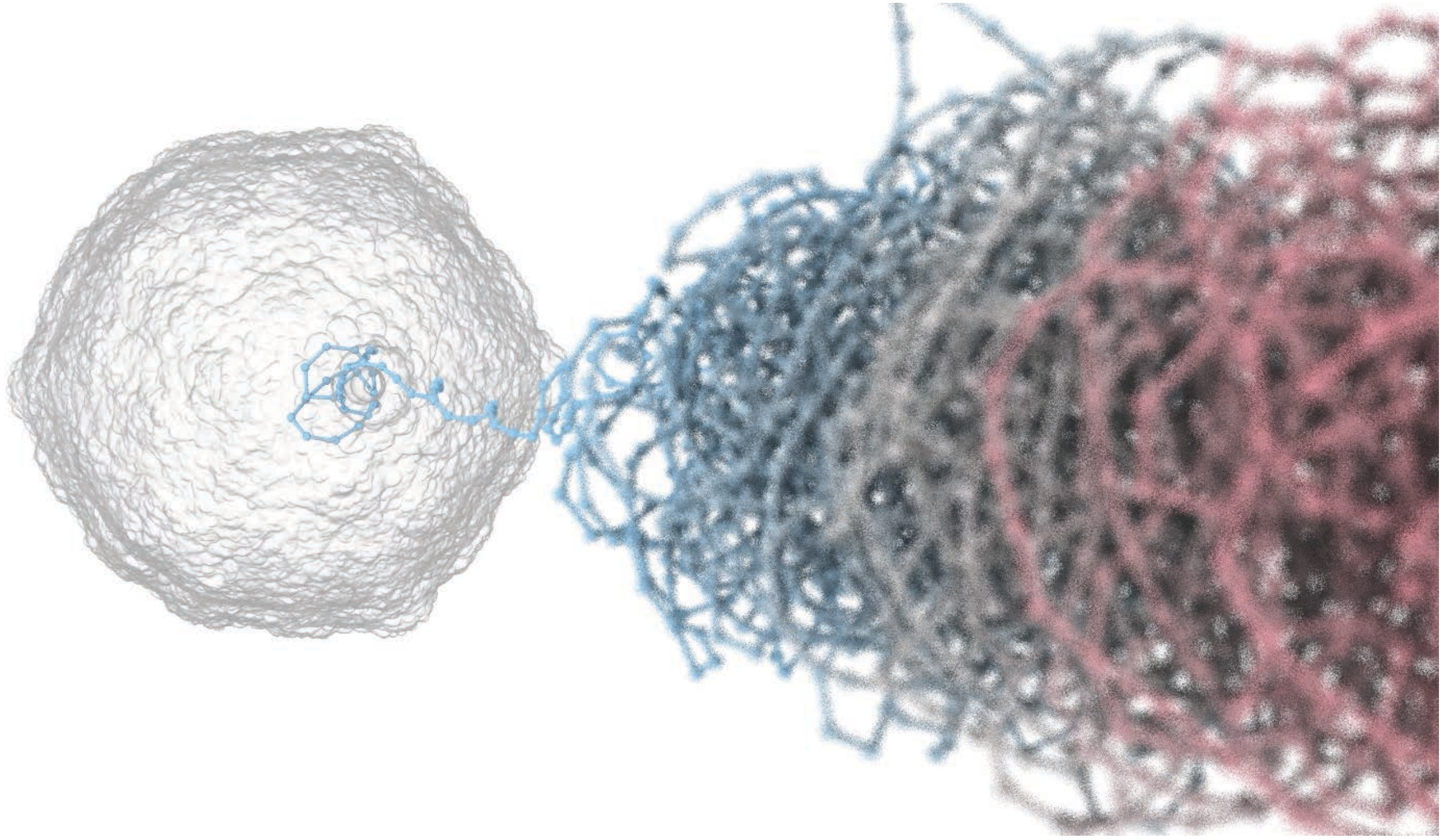
ARBD: Atomic Resolution Brownian Dynamics (multi-resolution)

Package DNA (CG) with ARBD, into CryoEM reconstruction of a HK97 bacteriophage capsid. A cryoEM map of the portal is fitted into the original capsid reconstruction, and DNA is packaged through the portal.



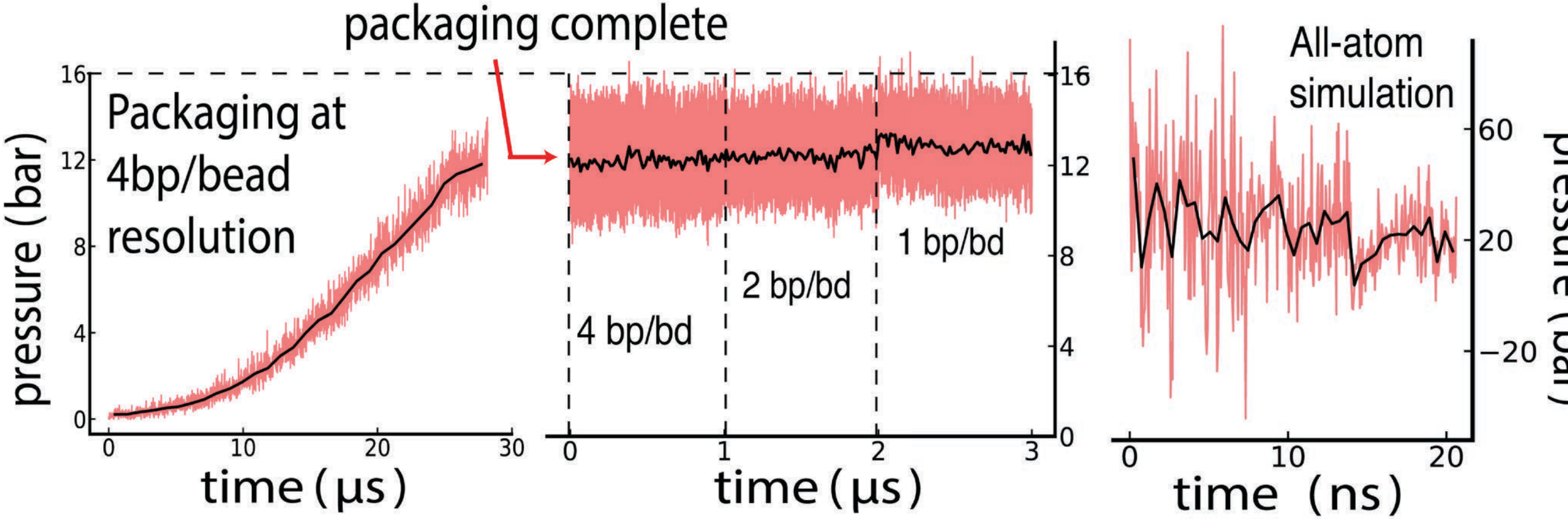
Smooth, purely repulsive grid-based potential obtained by blurring cryoEM density and adding the portal

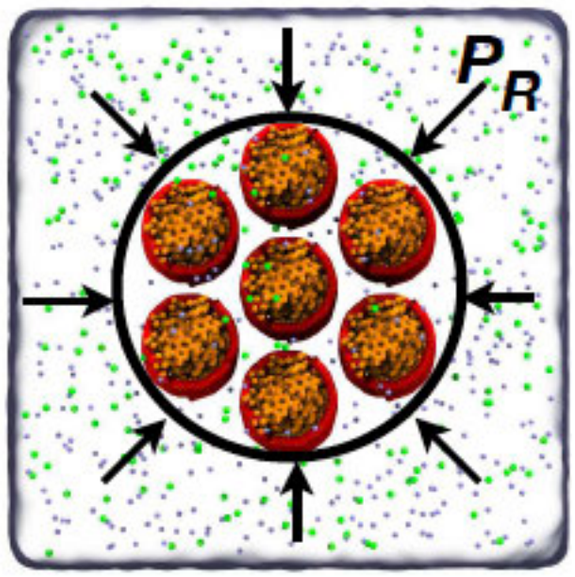
Multi-resolution packaging dsDNA viruses

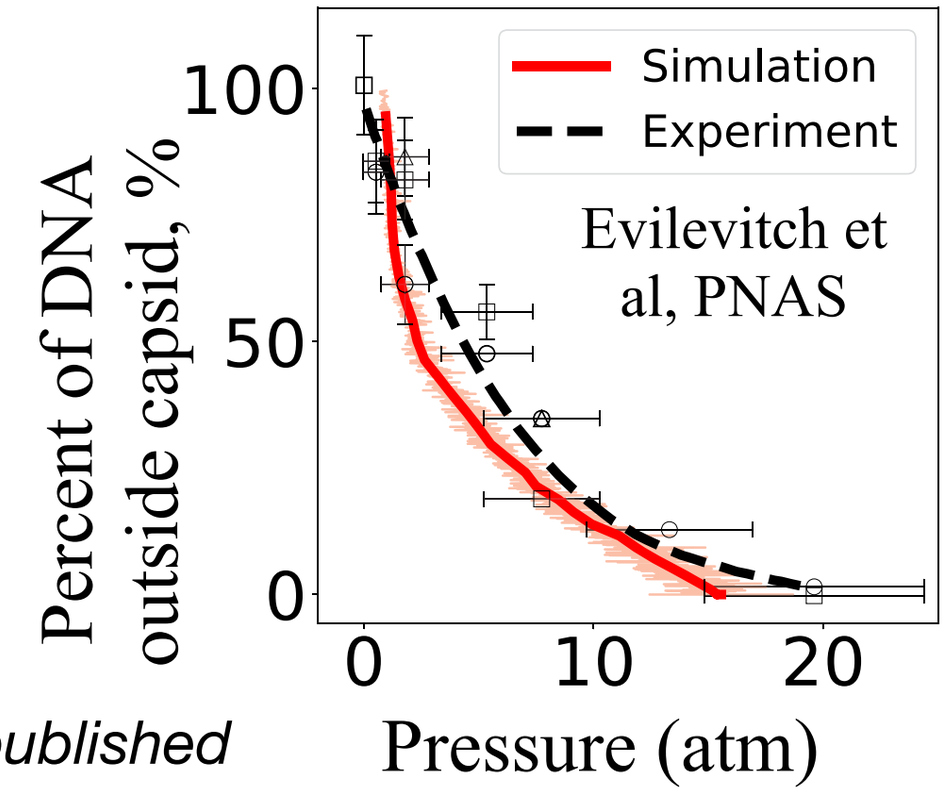


Kush Coshic et al. to be published

Internal pressure during packaging



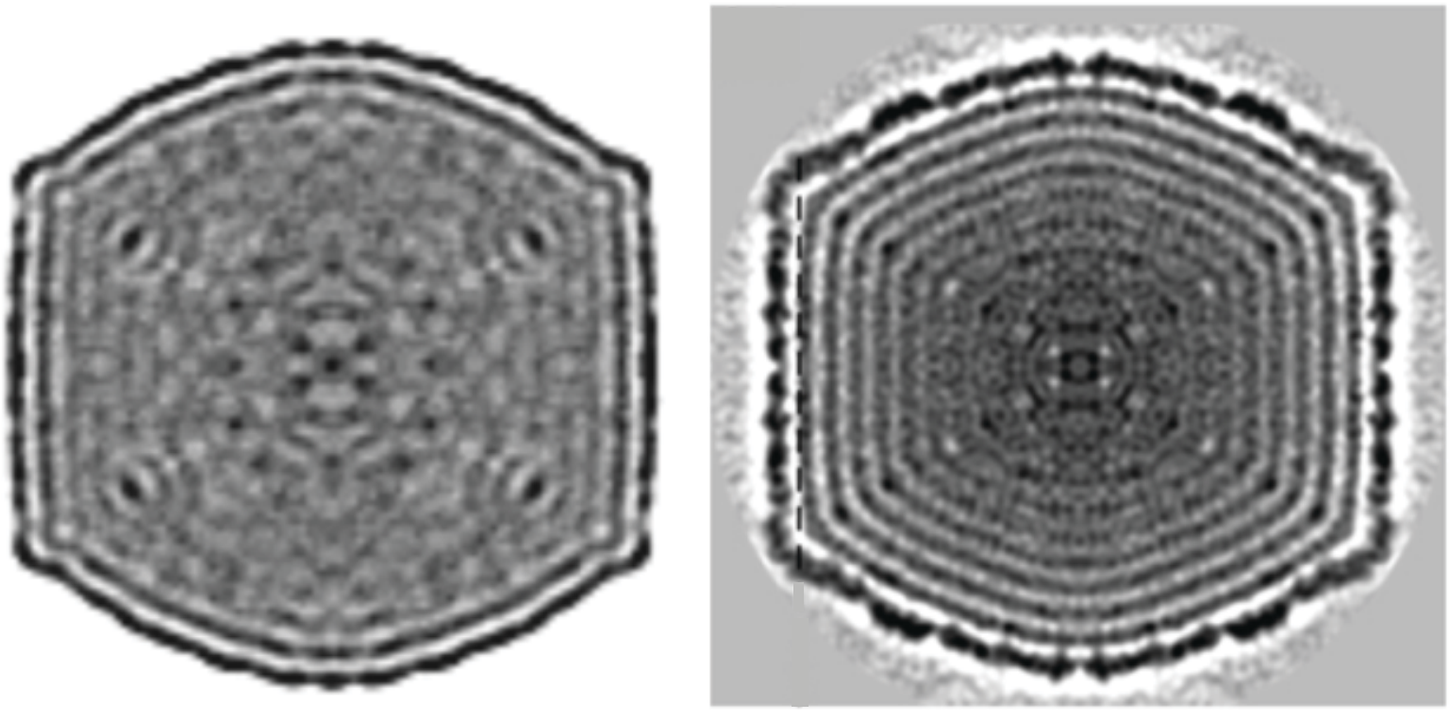




Kush Coshic et al. to be published

Comparison to structural data

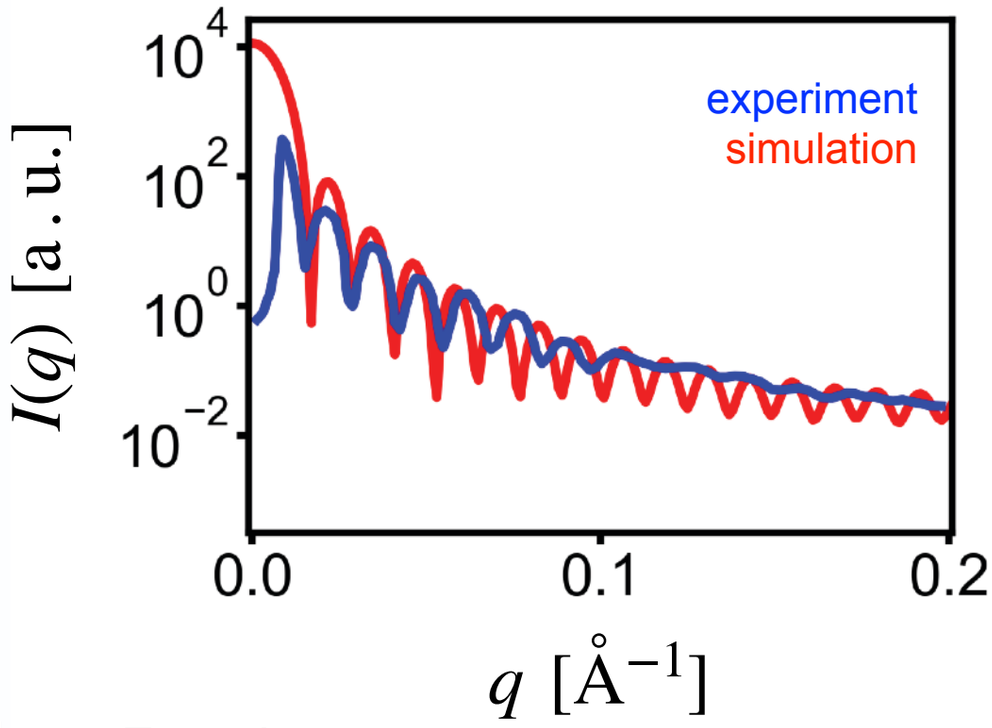
Cryo-electron microscopy



Simulation Experiment

J. Mol. Biol. (2009) 391, 471-483, Hendrix et al

Small Angle X-ray Scattering



Experiment:

Journal of molecular biology, 408: 541 (2011)

Simulation SAXS data were generated from CRYSOL, using an atomistic PDB of the protein coat and packaged DNA

Kush Coshic et al. to be published

Conclusions and outlook

The length and time scale of an all-atom MD allows for adequate sampling on DNA-DNA interactions in complex environment

All-atom force field is accurate enough to make quantitative, experimentally testable predictions

A multi-resolution representation can expand the time and length scale of processes amenable to all-atom MD approach

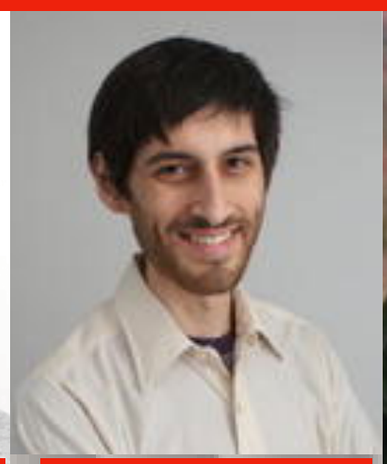
Acknowledgements

- Funding through CPLC

- Computations

UIUC team

Jejoong Yoo



Chris Maffeo

Kush Coshic

David Winogradoff

TJ Ha JHU

Taekjip Ha

Kim group at UNIST in Korea

Hajin Kim

Hyunju Kang