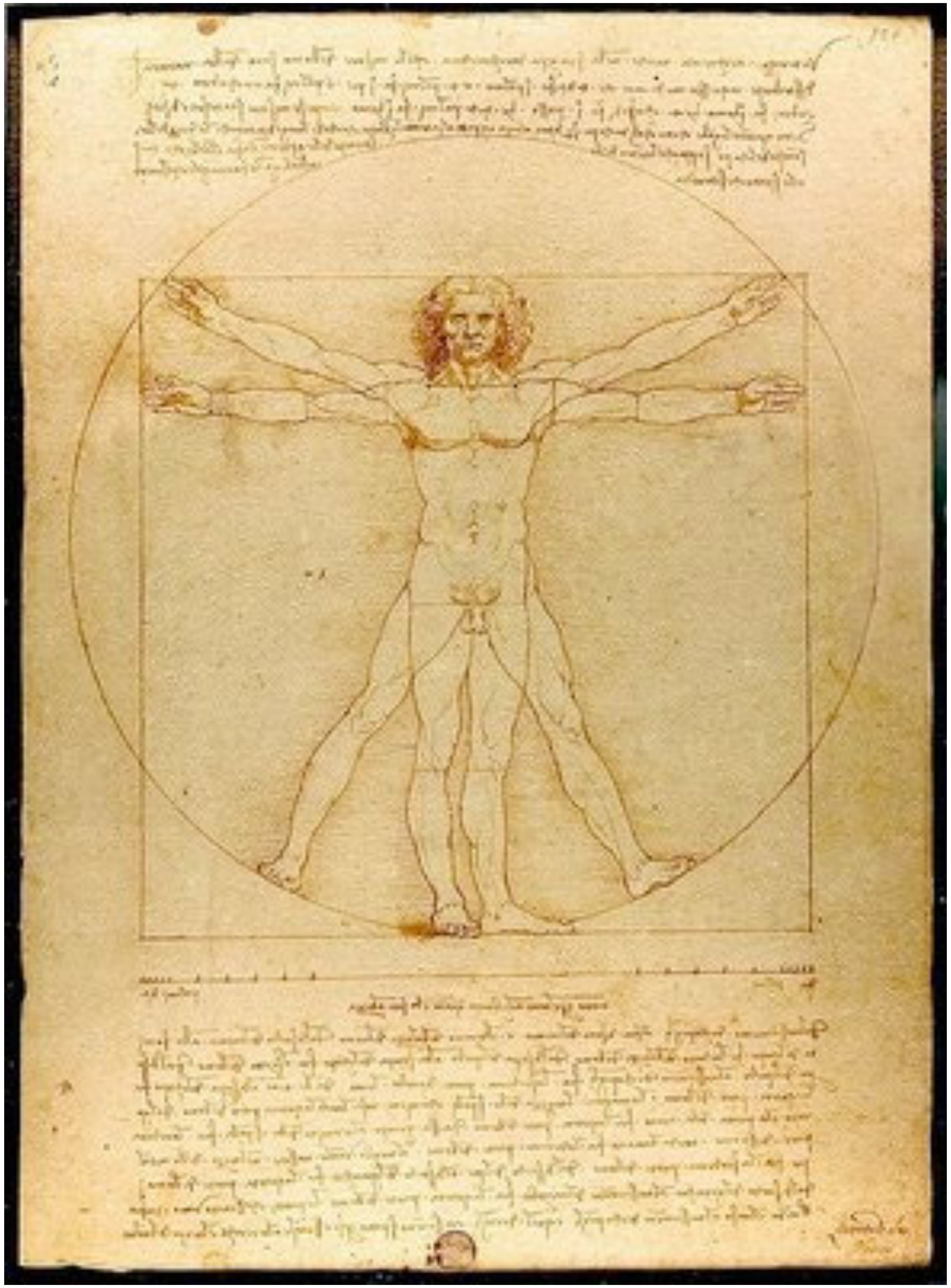
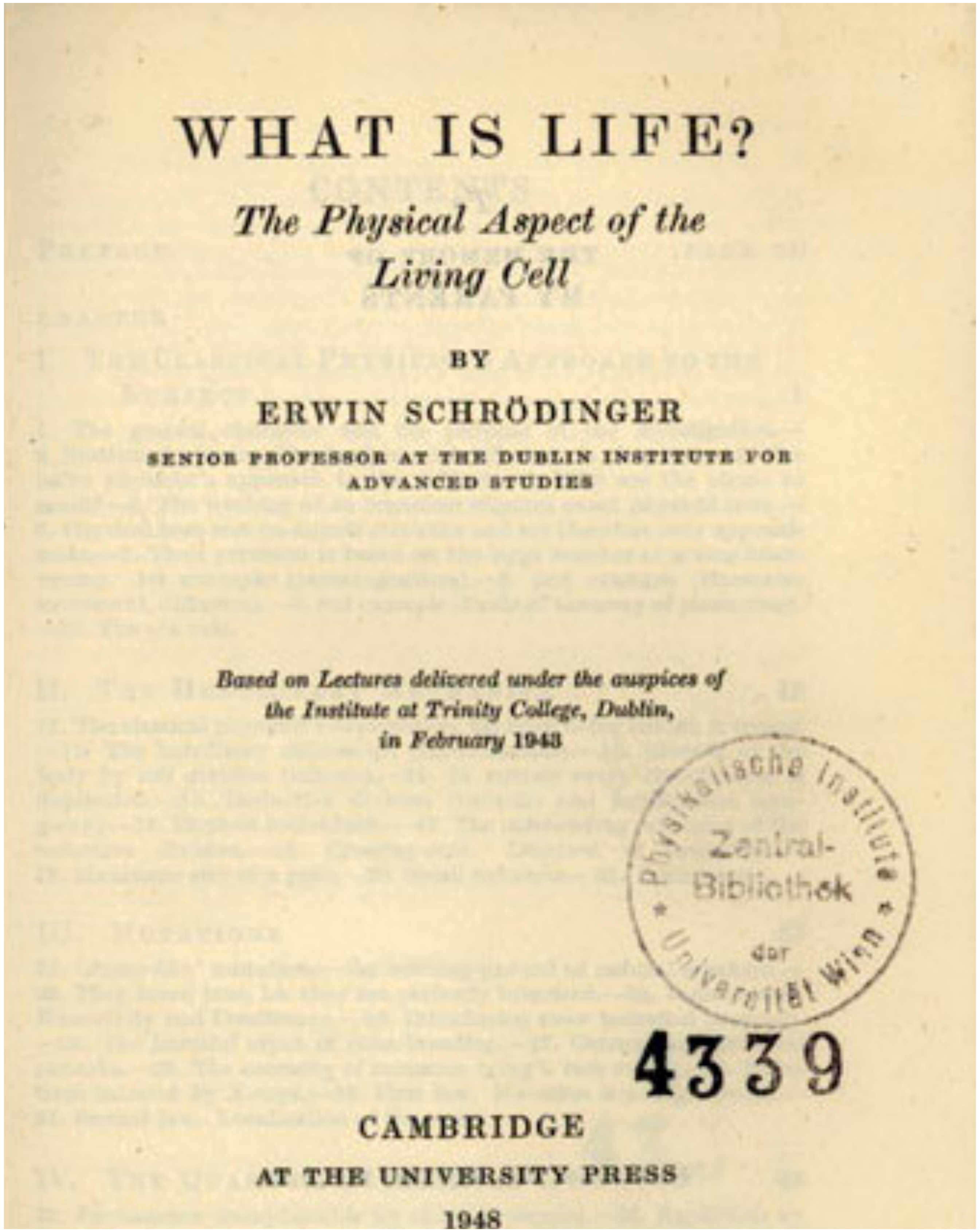
Introduction to Simulation and Modeling of DNA Systems

Aleksei Aksimentiev
Department of Physics, University of Illinois at Urbana -Champaign





WHAT IS LIFE?

The Physical Aspect of the Living Cell

BY

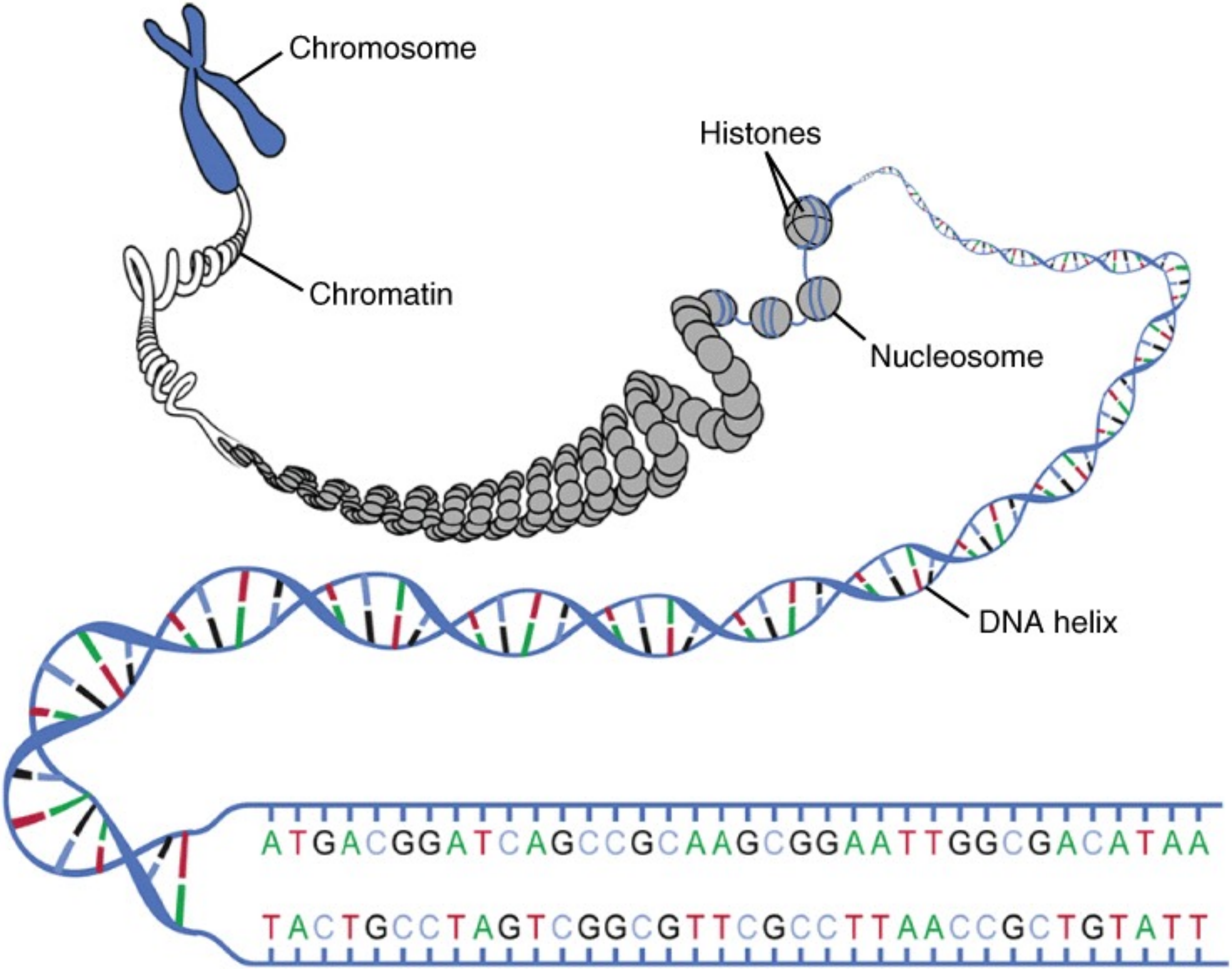
ERWIN SCHRÖDINGER

SENIOR PROFESSOR AT THE DUBLIN INSTITUTE FOR ADVANCED STUDIES

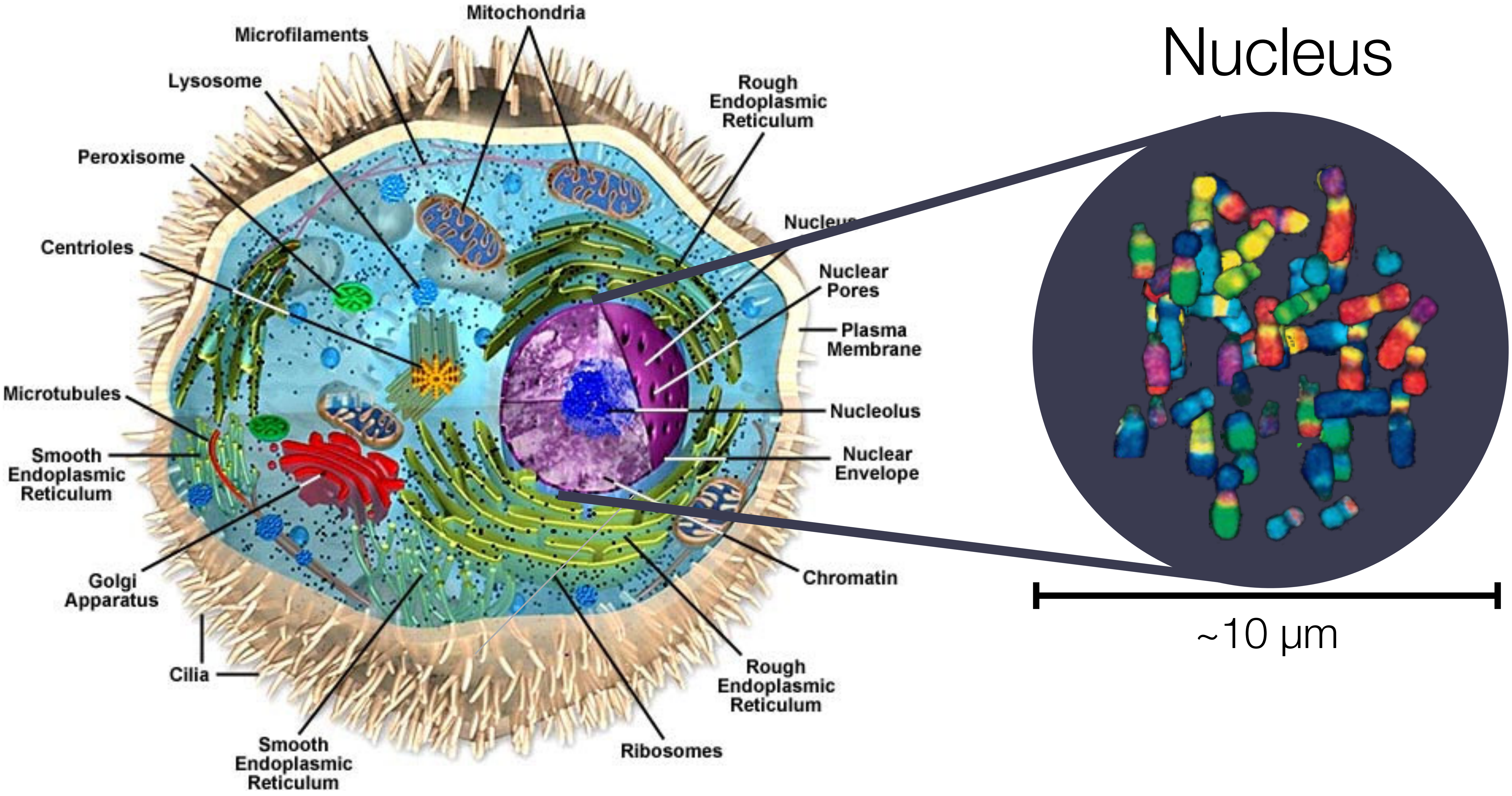
Based on Lectures delivered under the auspices of the Institute at Trinity College, Dublin, in February 1943

CAMBRIDGE
AT THE UNIVERSITY PRESS
1948

DNA, the blueprint

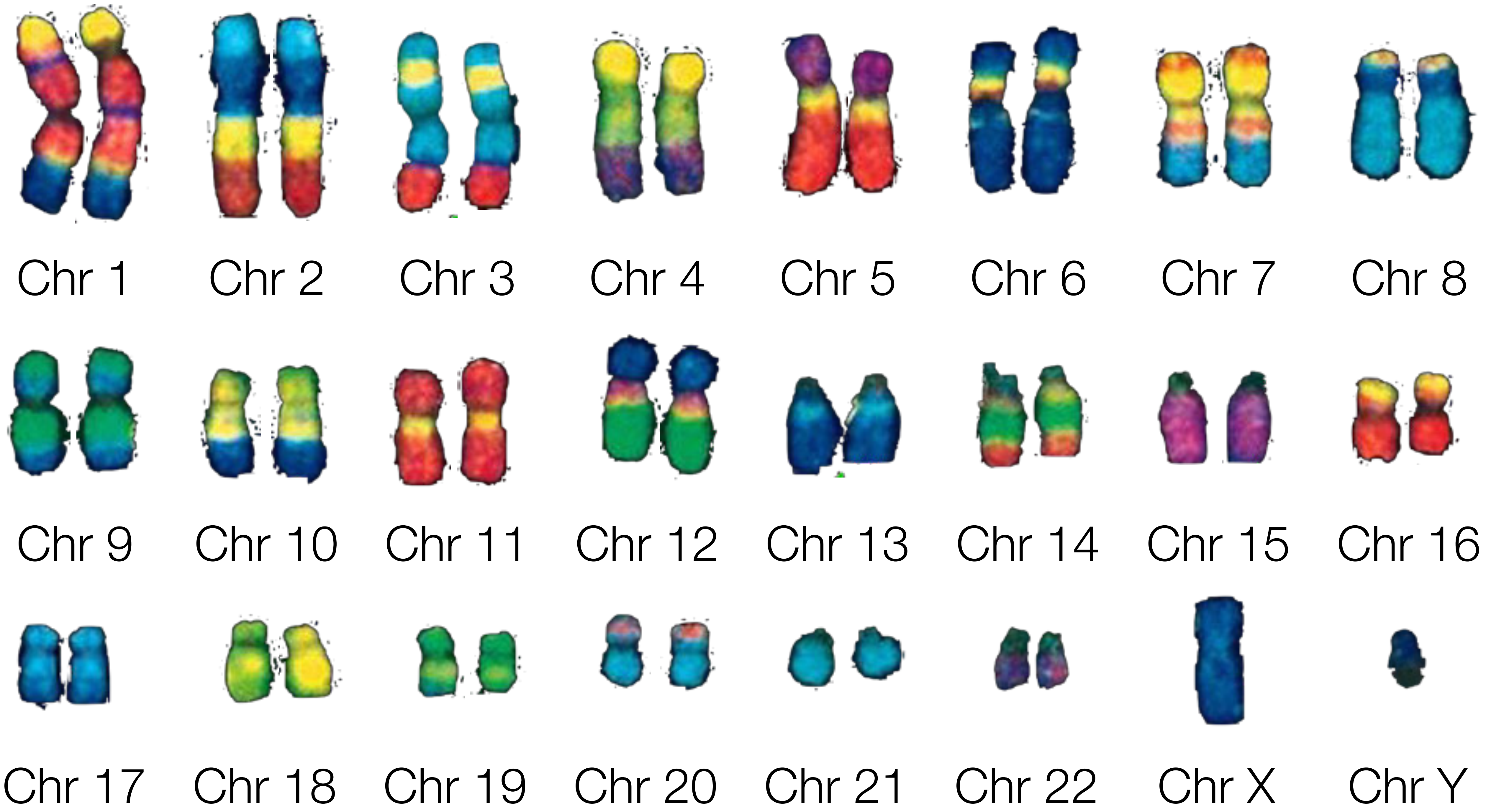


A nucleus of a human cell contains 23x2 chromosomes



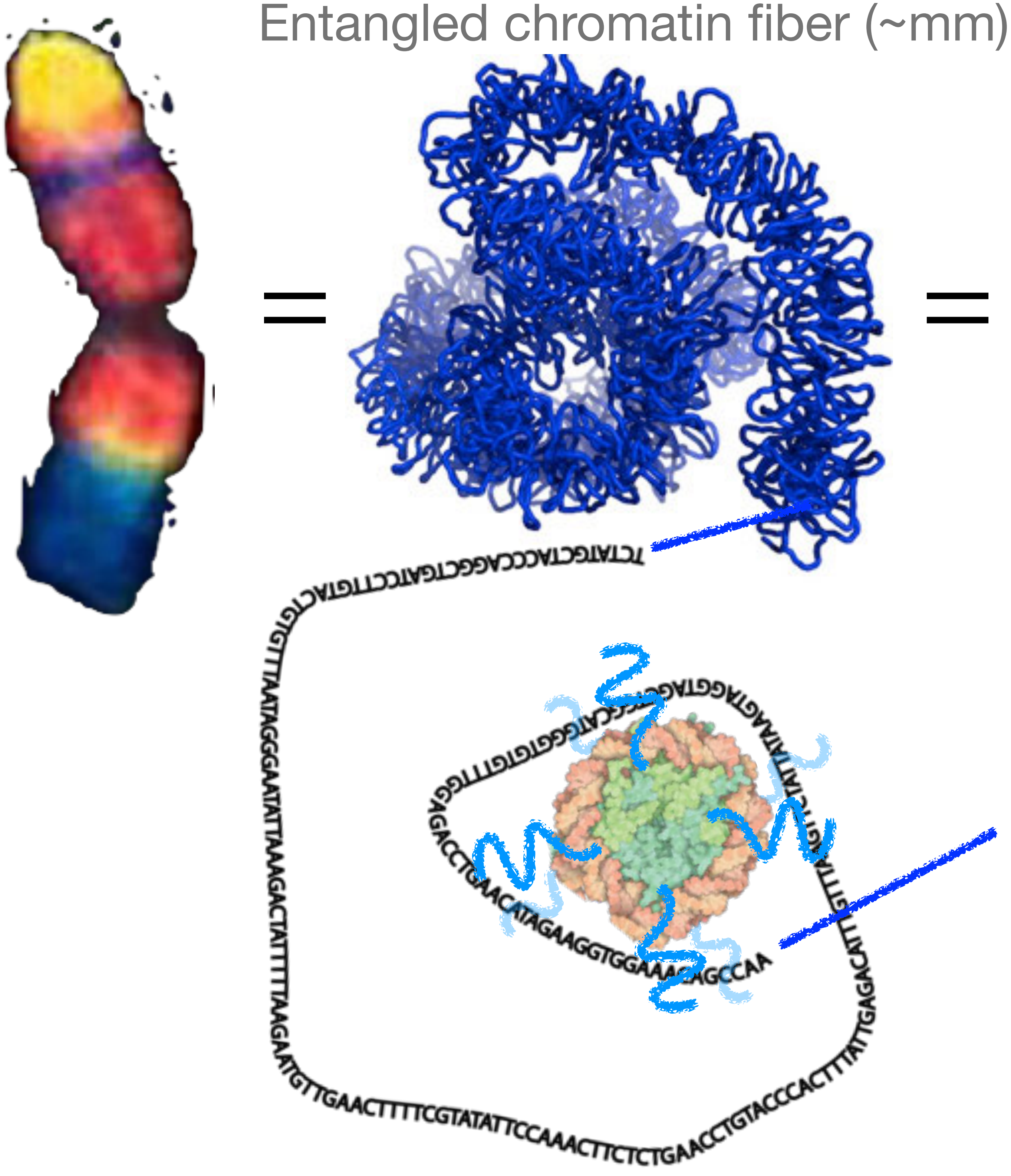
http://homepage.smc.edu/wissmann_paul/histology/

Human chromosomes



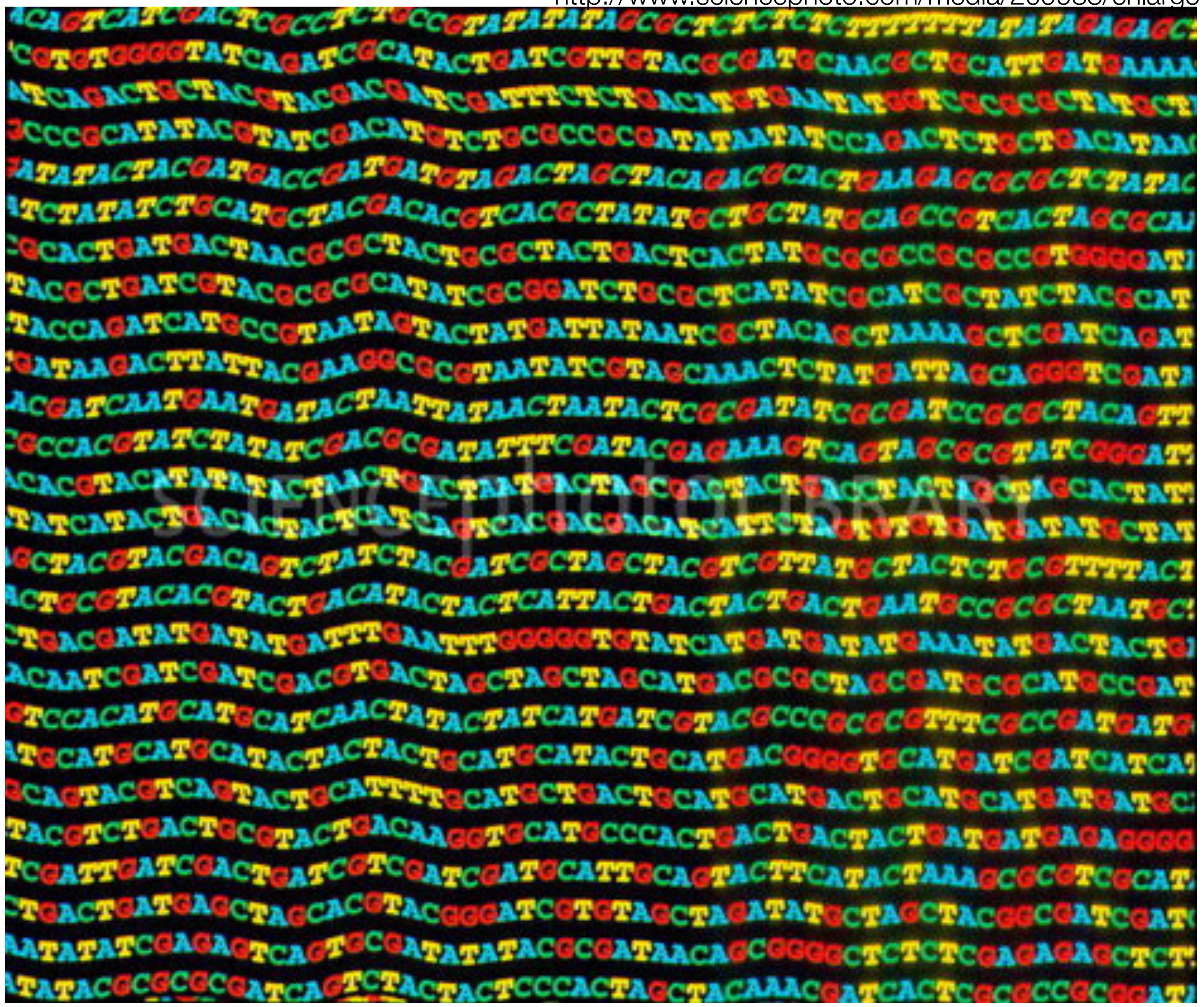
Each chromosome is millimeters-long single fiber: chromatin

Entangled chromatin fiber (~mm)



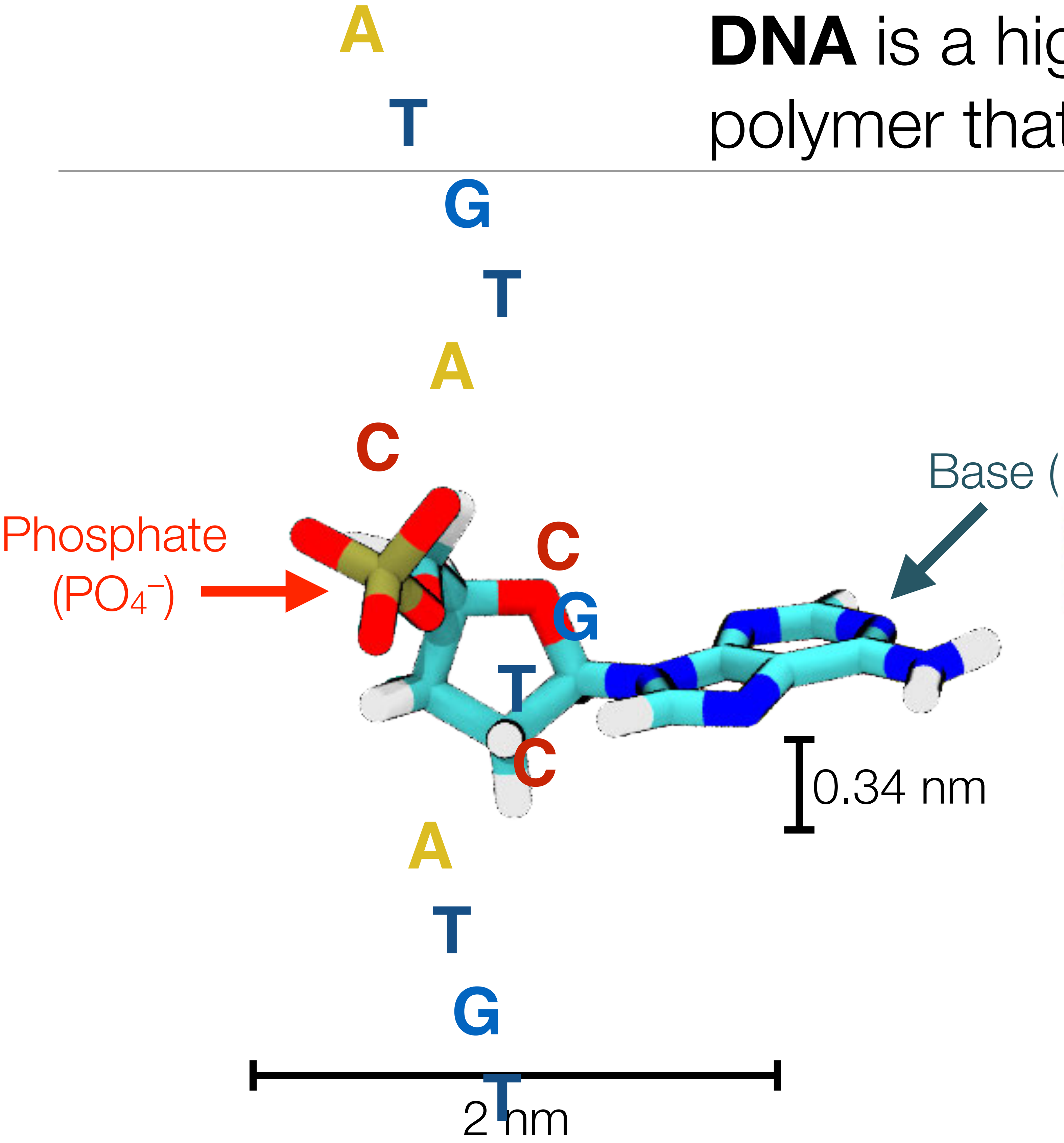
=

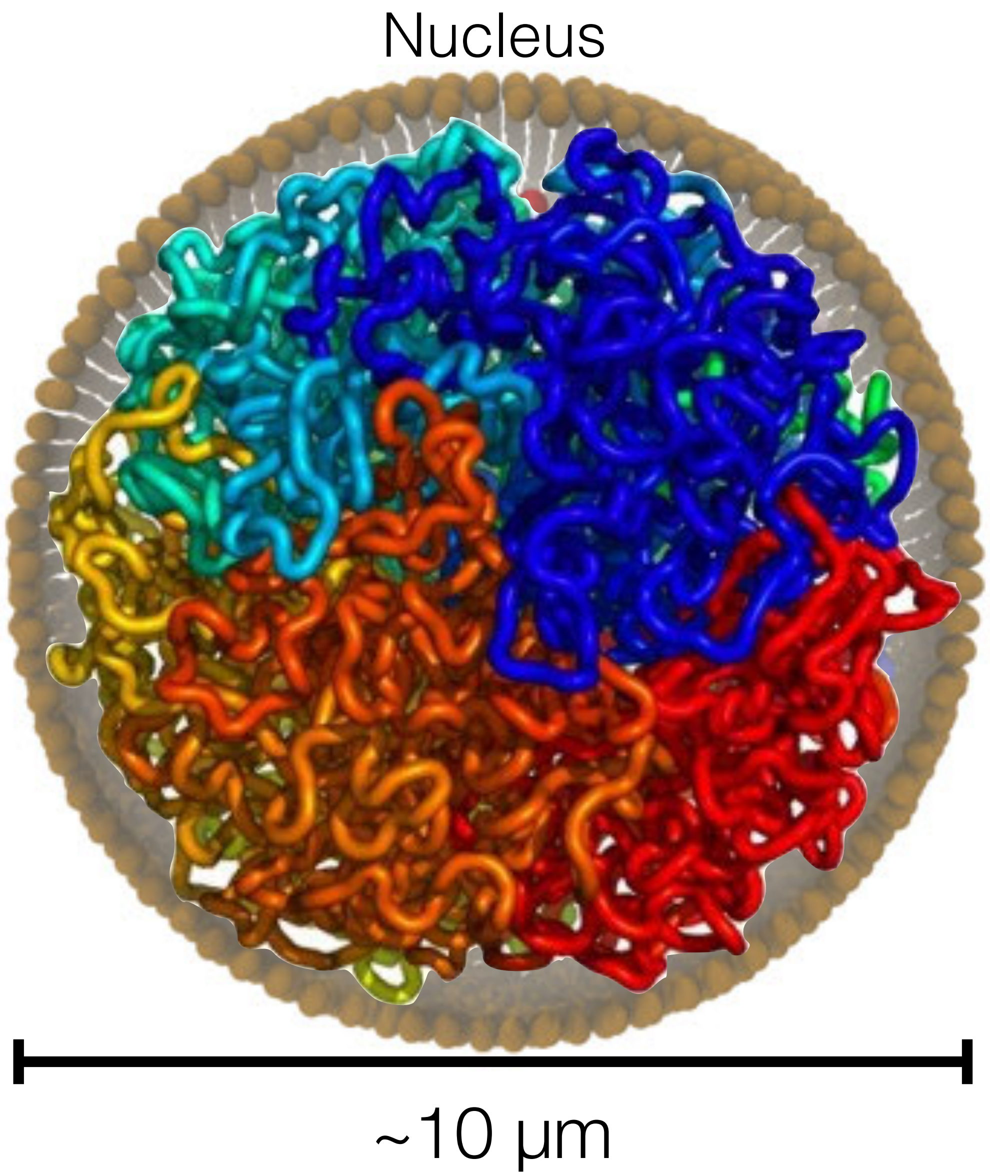
http://www.sciencephoto.com/media/209983/enlarge



Human genome:
3 billions of A, T, G and C

DNA is a highly *negatively charged* polymer that forms human genome





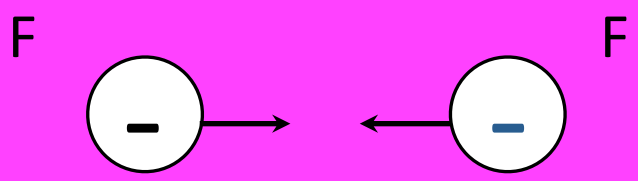
A human nucleus contains 23x2 chromosomes made of 12 billion nucleotides (~2 meters of DNA)

Animation by Chris Maffeo

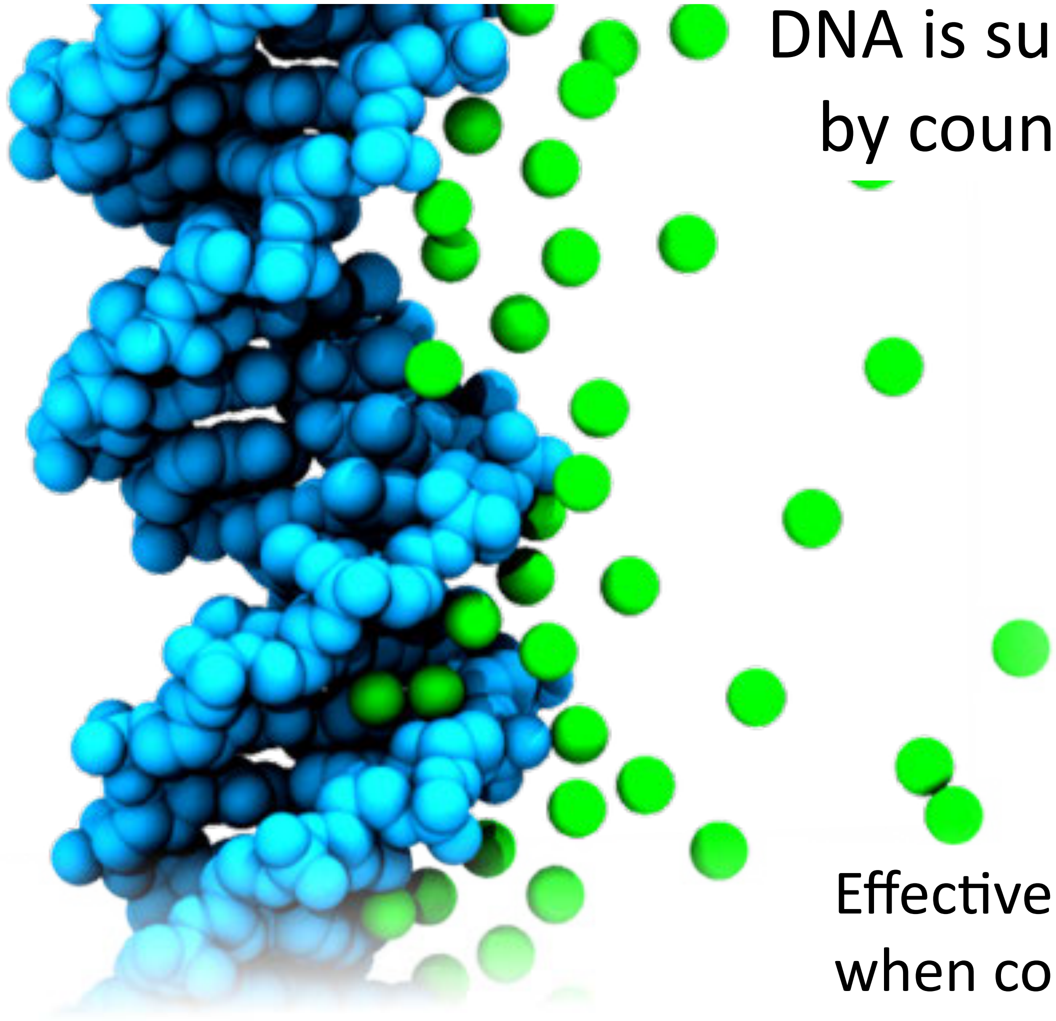
Lieberman-Aiden et al., Science (2009)

Same sign charges ….

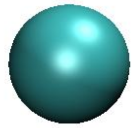
Same sign charges repel
(in vacuum)



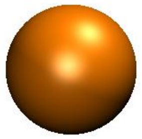
Same sign charges can
attract (in a medium)



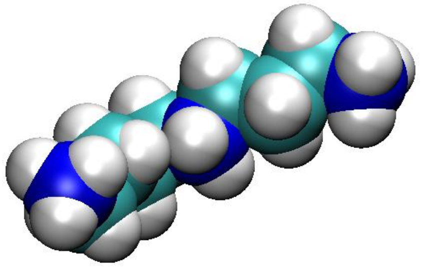
DNA is surrounded
by counter ions



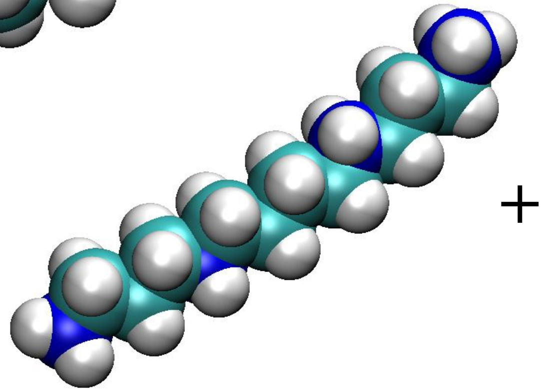
+1e, sodium or potassium



+2e, magnesium or calcium



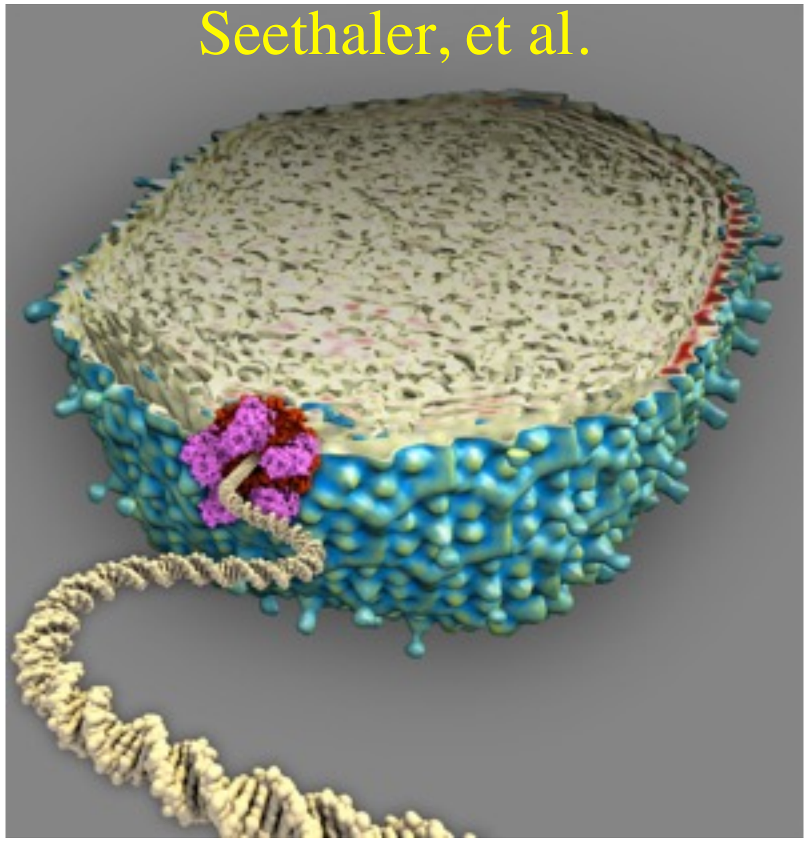
+3e,spermidine



+4e,spermine

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

MD simulation of dense DNA arrays


Seethaler, et al.

What we control

DNA density (or harmonic constraint radius)

$[Na^+]_{buf}$ ~ 200 mM

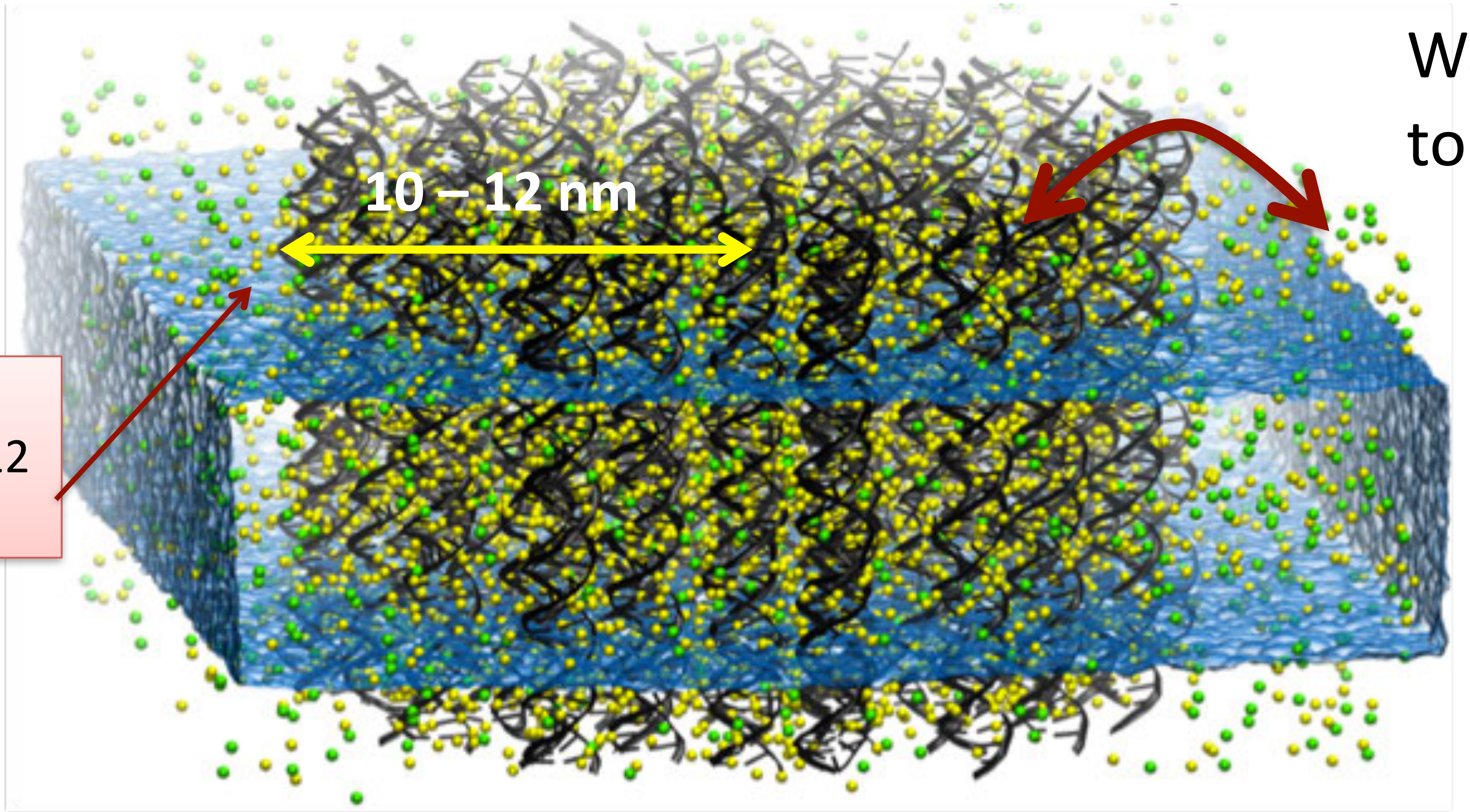
$[Mg^{2+}]_{buf}$ ~ 0 or 20 mM

What we measure

Pressure as a function of [ion] & [DNA]

DNA / ion distribution:

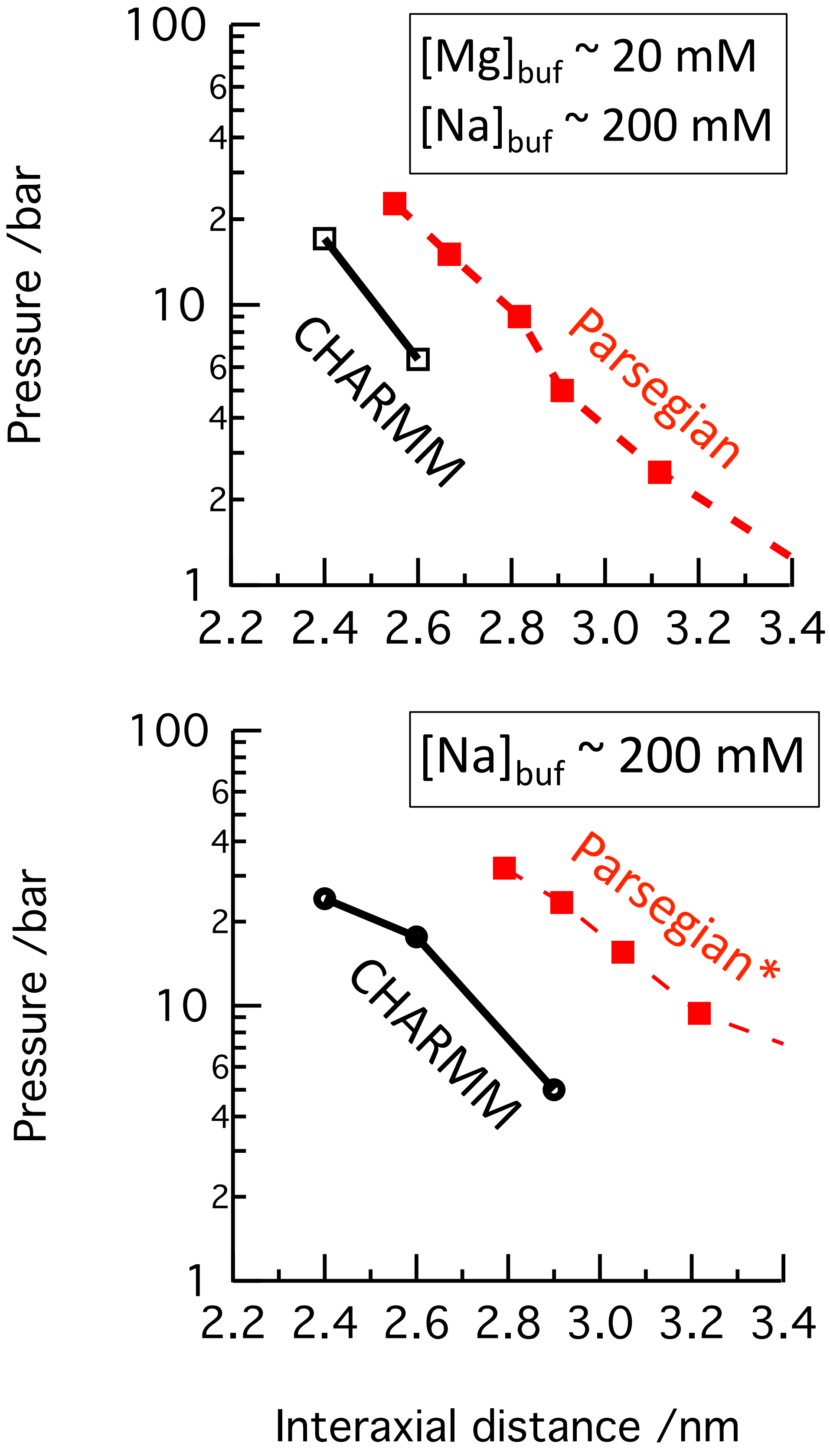
DNA / ion diffusion inside the array



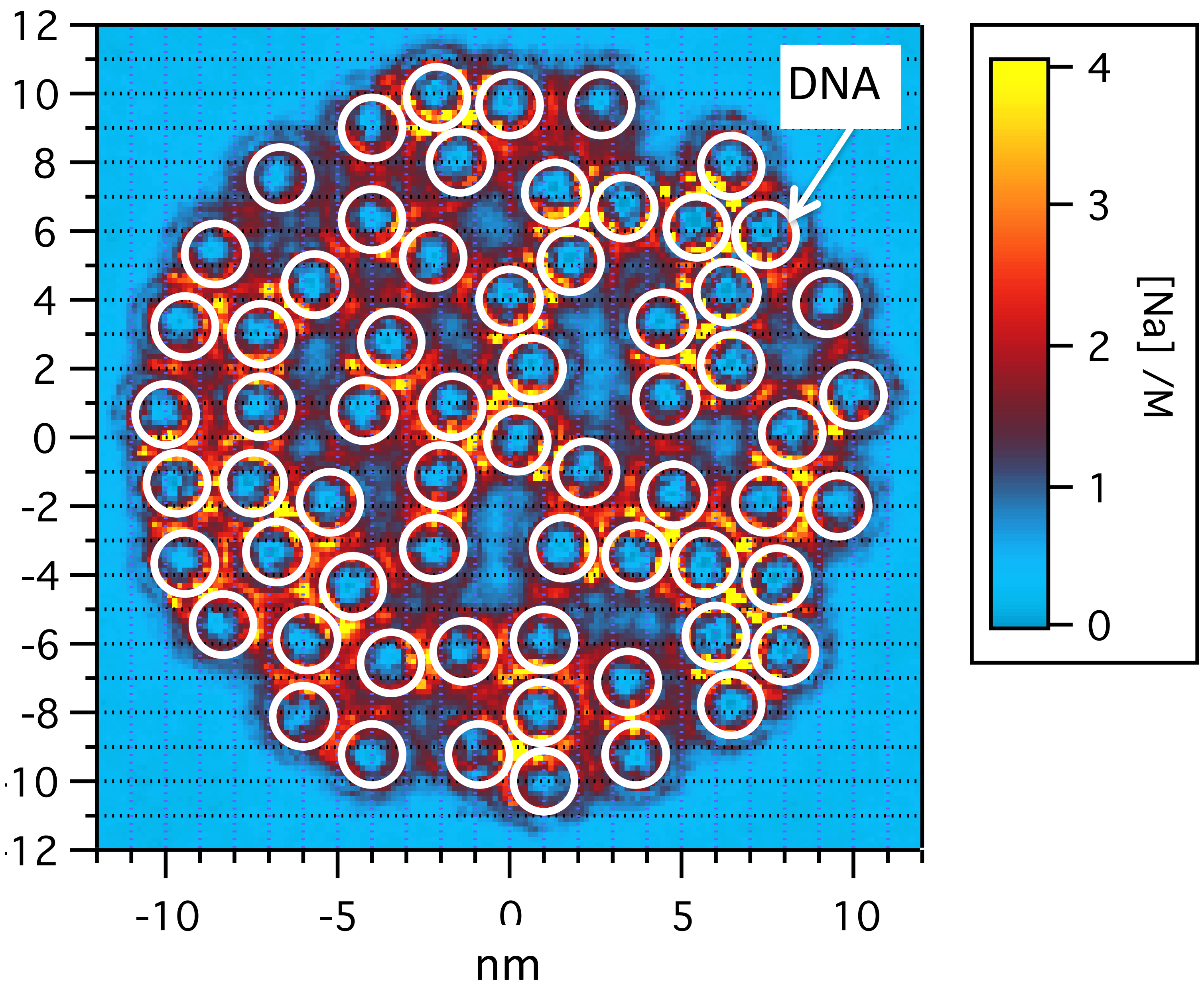
Water / salt free to move

Cylindrical harmonic constraint (radius of 10 – 12 nm) only against DNA

The standard MD force field fails to predict internal pressure of a DNA array

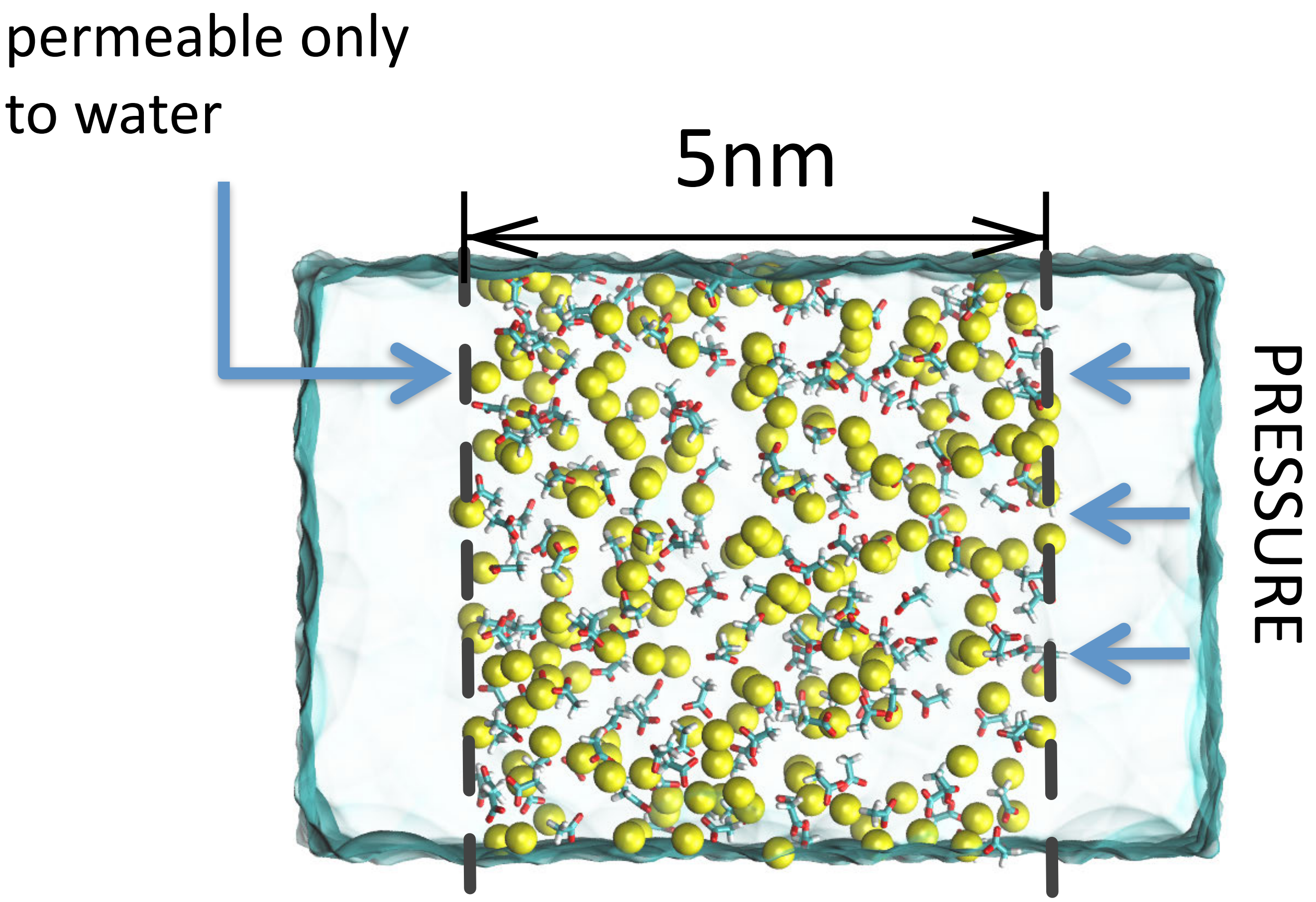


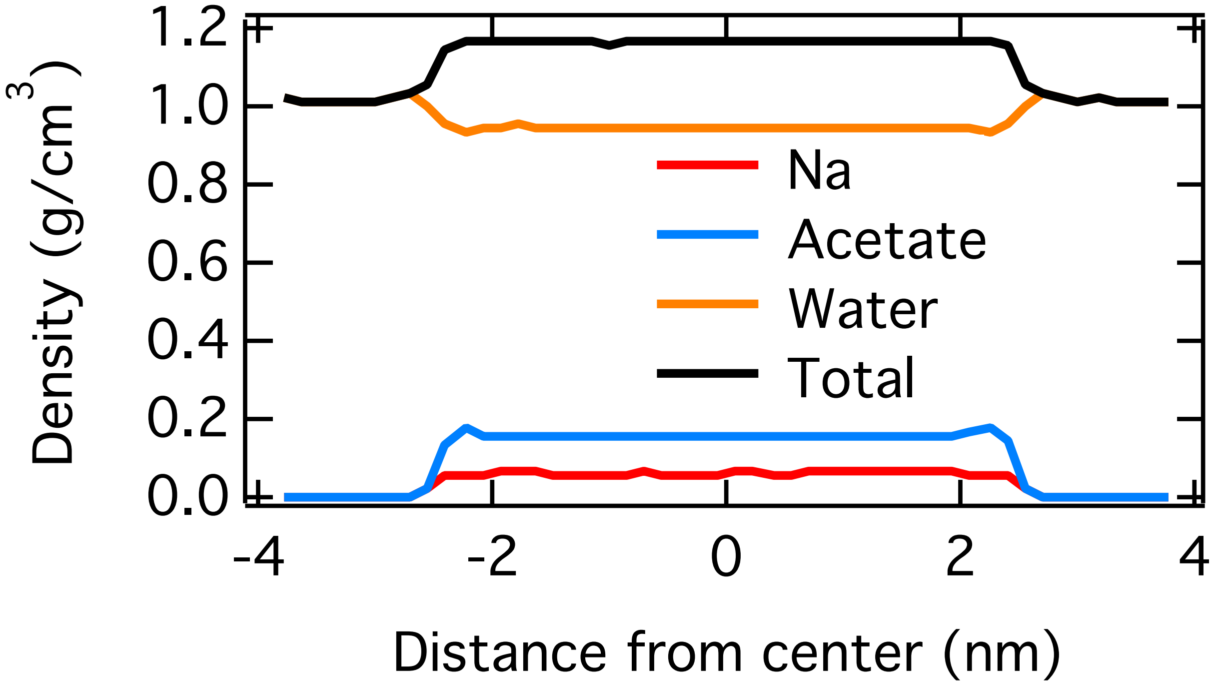
Too strong Na/Mg-phosphate attraction induces artificial DNA clusters!! [Na] ~ 4M!!



* Rau, D. C.; Lee, B.; Parsegian, PNAS (1984)

Recalibrate ion-DNA parameters using osmotic pressure data



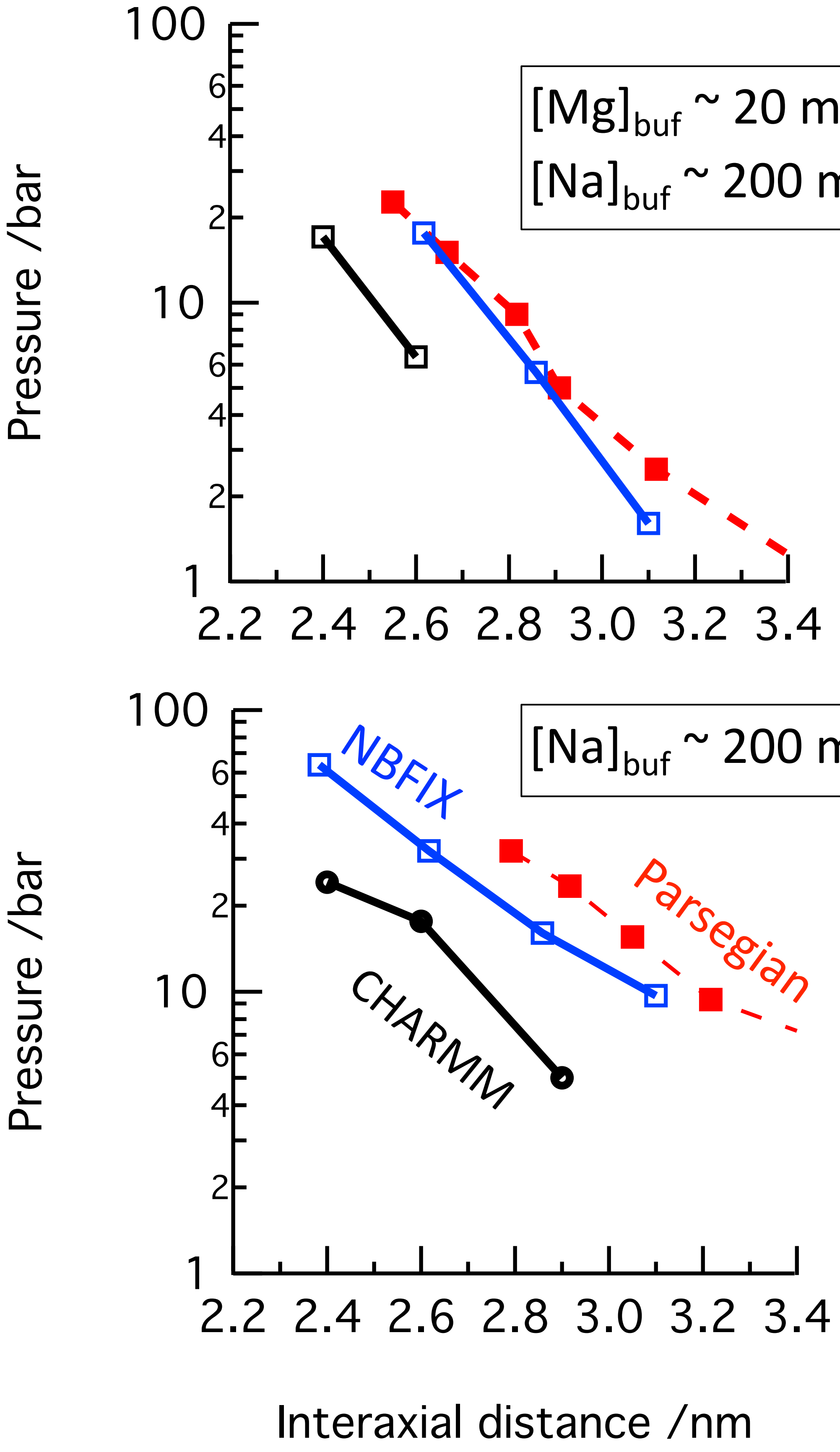


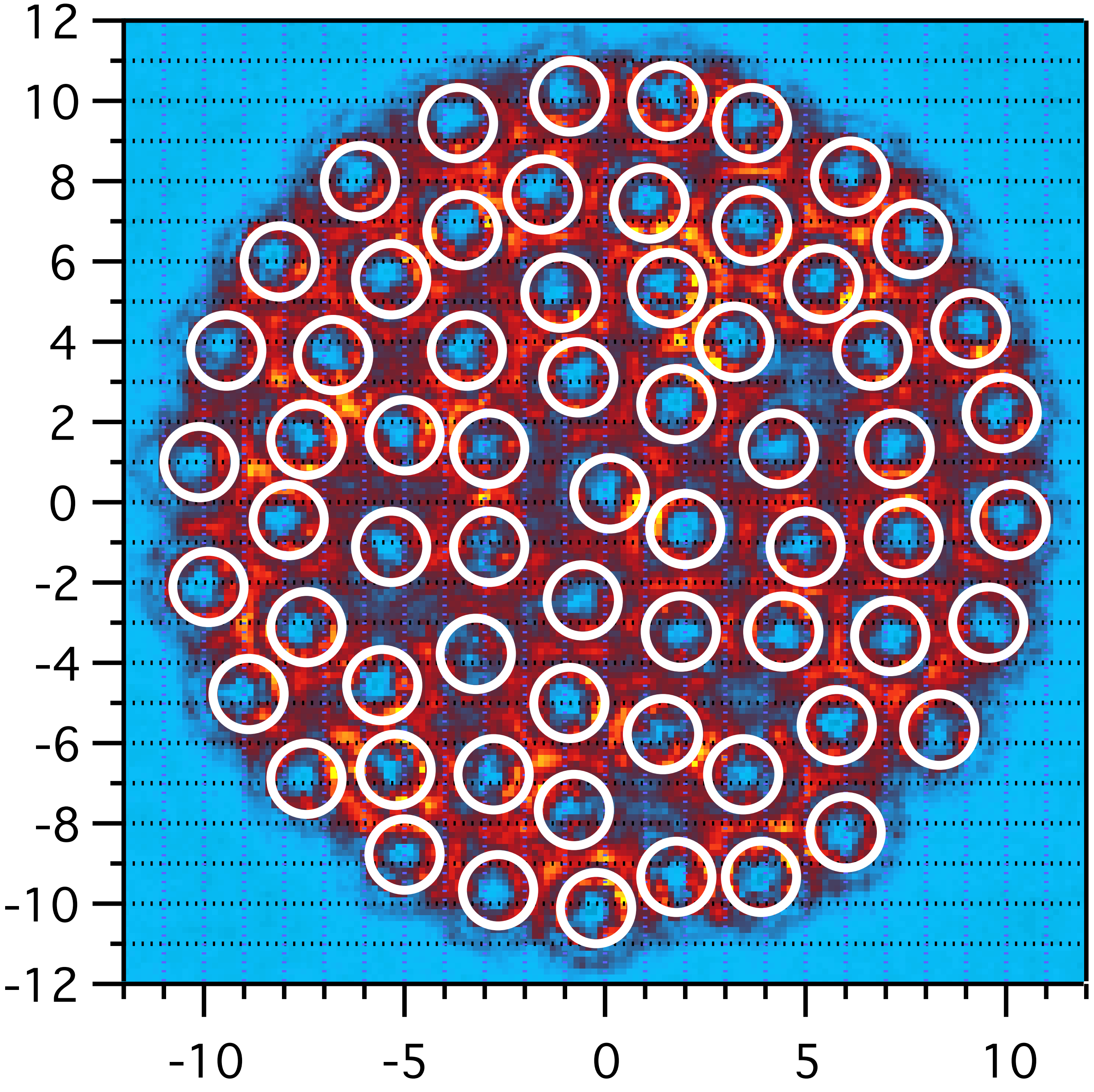
- Osmotic pressure is directly related to ion-pair formation: $\pi = \phi c R T$

- Pros: modify only ion-DNA phosphate interaction, without altering ion-water interaction.
- Cons: nothing.

* Luo & Roux, JPCL (2009)

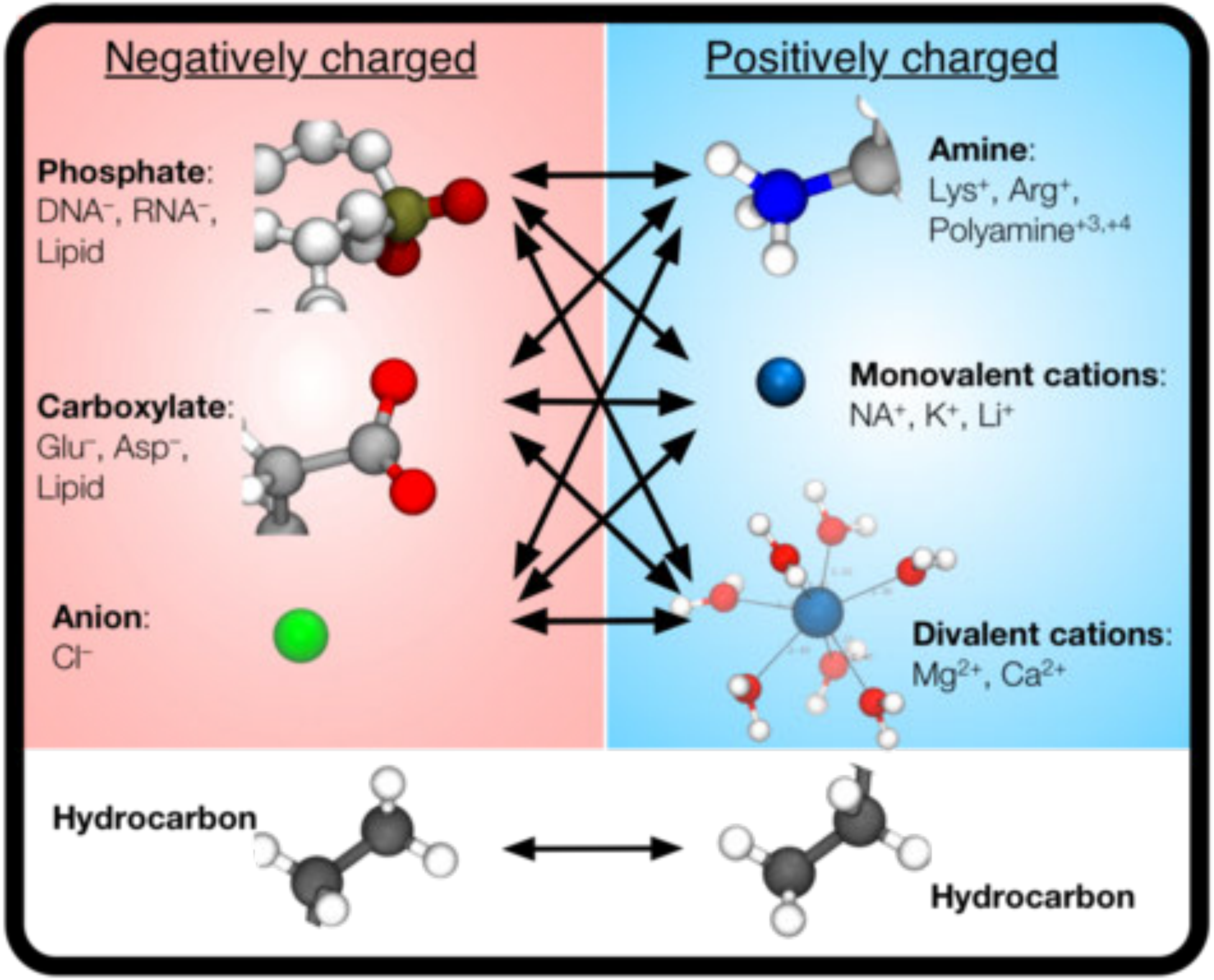
Improved parametrization of ion-DNA interactions



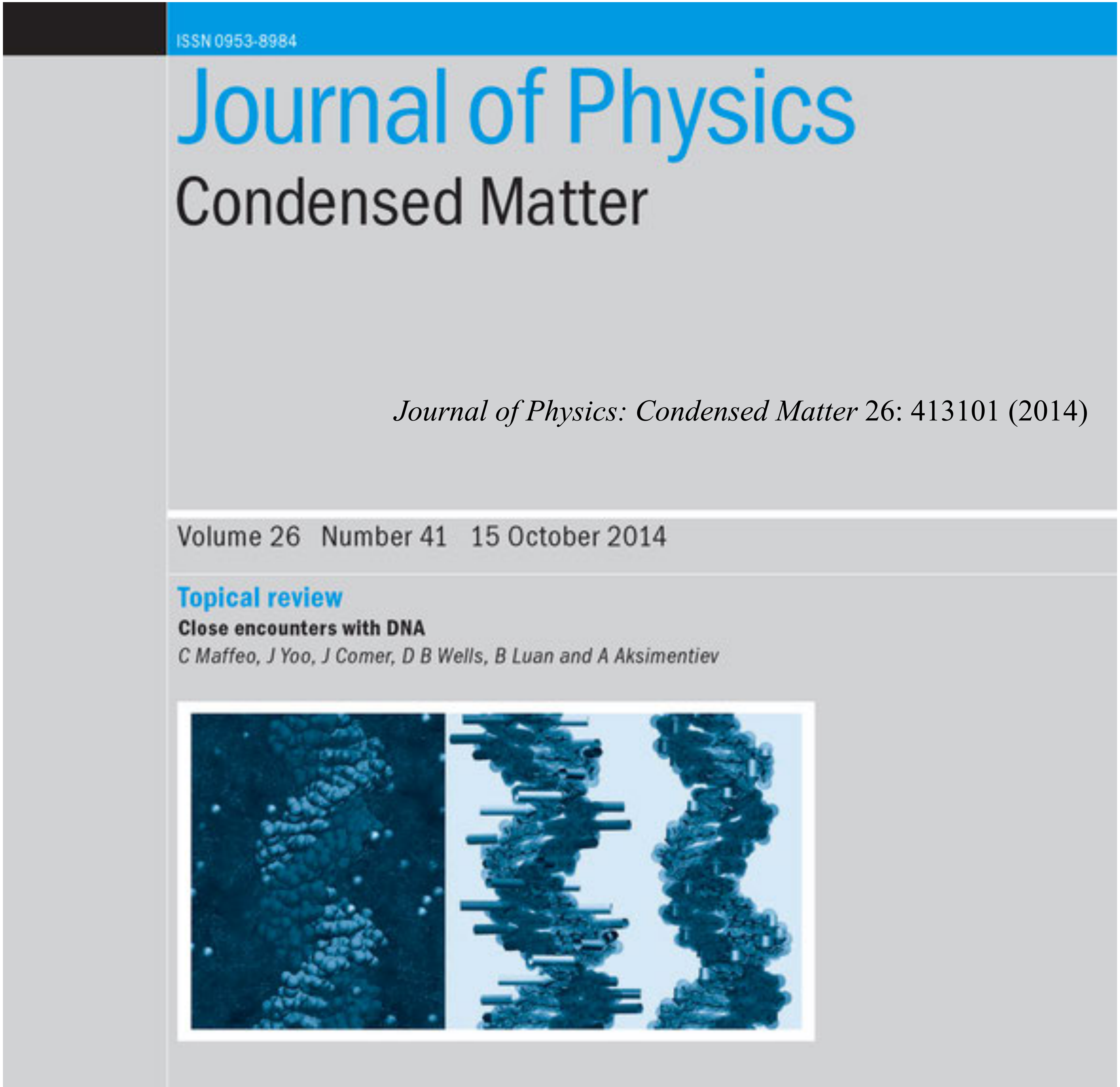


J. Phys. Chem. Lett. 3:45 (2012)

CUFIX: Accurate parameterization of non-bonded interactions



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



ISSN 0953-8984

Journal of Physics
Condensed Matter

Journal of Physics: Condensed Matter 26: 413101 (2014)

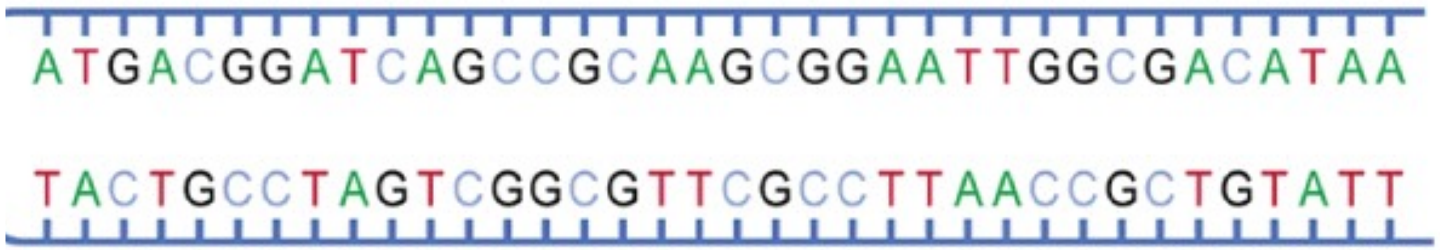
Volume 26 Number 41 15 October 2014

Topical review

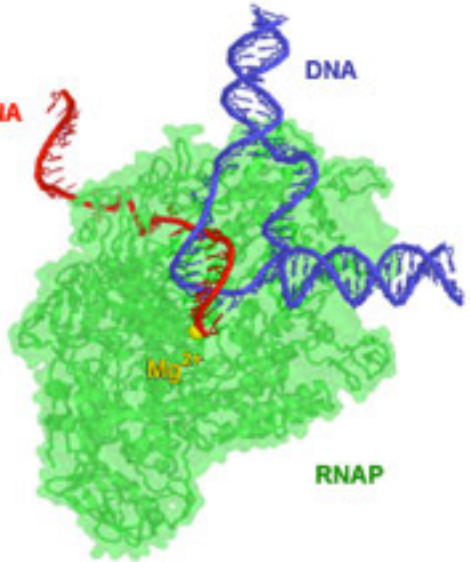
Close encounters with DNA

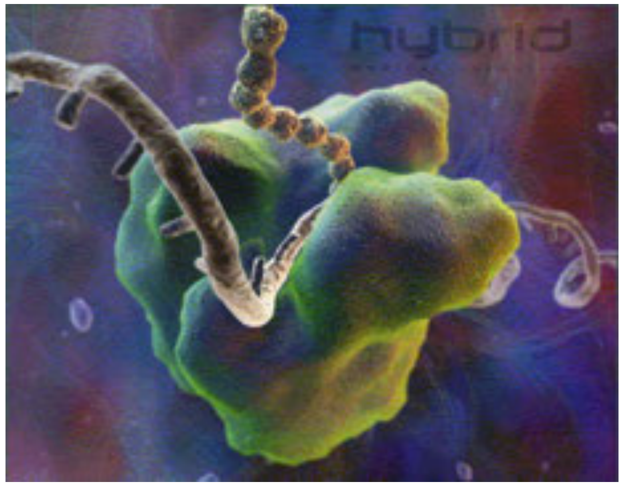
C Maffeo, J Yoo, J Comer, D B Wells, B Luan and A Aksimentiev

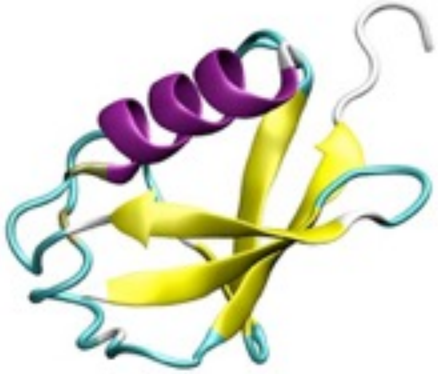
The nucleotide sequence contains biological information

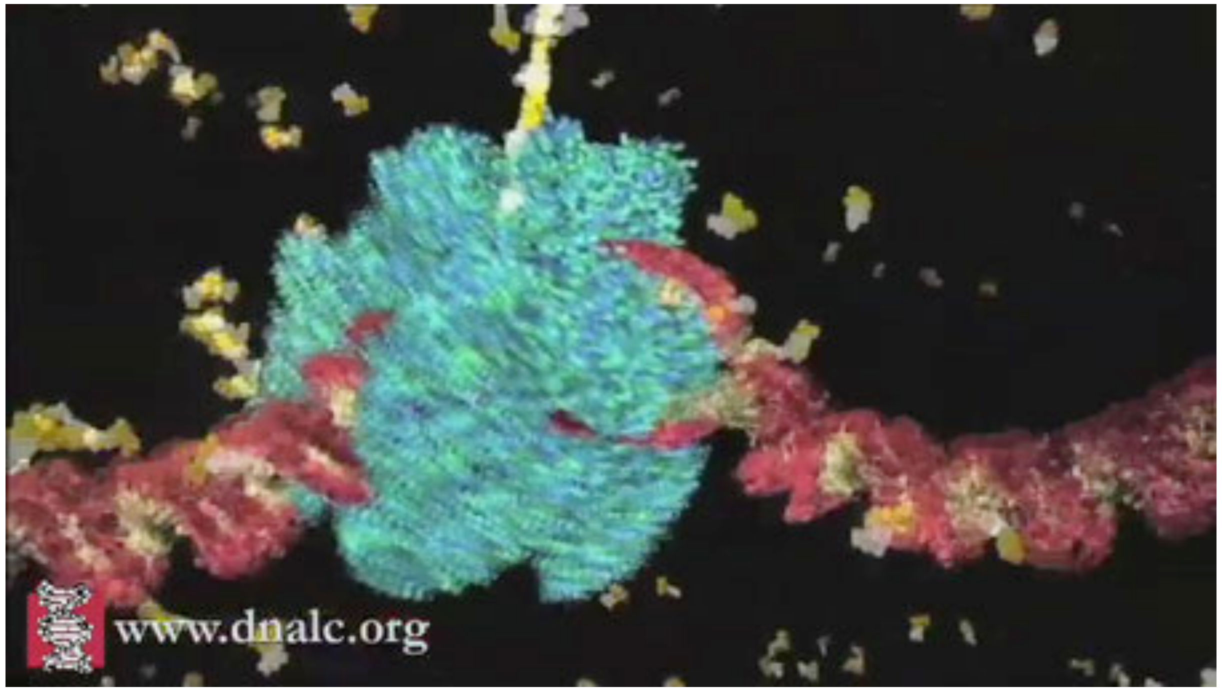


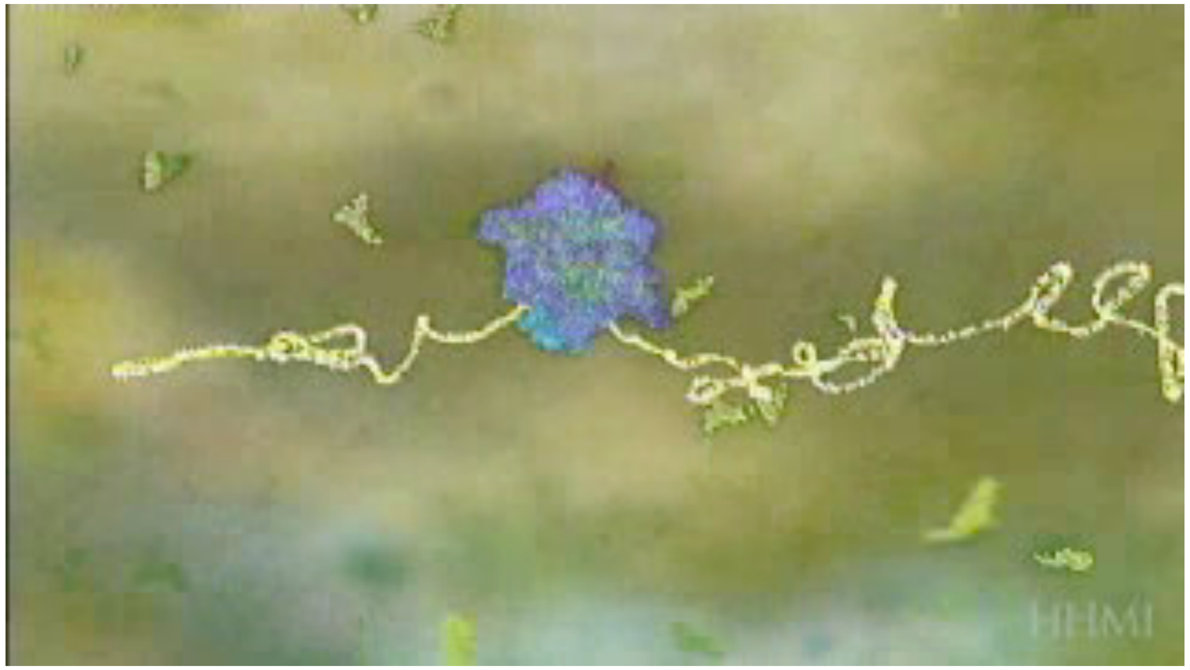
Central dogma of molecular biology







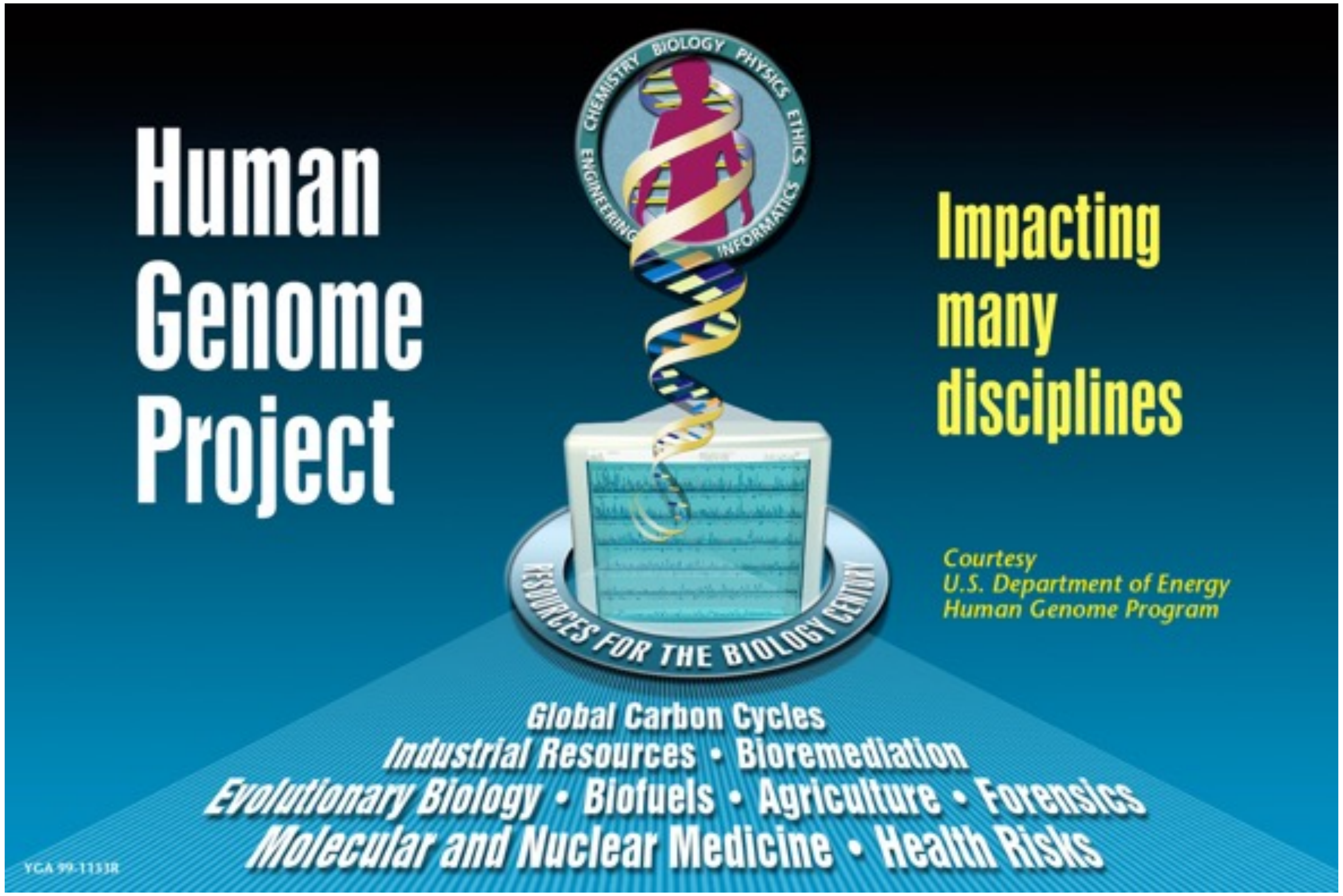




The Human Genome Project

Duration:
October 1990 - 2003

Discovered <u>ALL</u>
20,000-25,000
human genes



Courtesy
U.S. Department of Energy
Human Genome Program

Determined complete sequence of the 3 billion DNA bases

5’-ACCGGTGGGTGCATAGCTGTGCTGTAAGTGAAGTG
AGGCGGCAGGTGTTGAAAGTCGATGTAGTTCGTAG
GTCAGTTGATGTCGATGTGAAATGCTGATGCTAGTG
GACAGGGTGACTAGTGAATCGATGCTAGCCTAGCTA
GTCAGTGGTGCTAGCTACGATCGATTTCAGGCTGCT

GTGGGTGCATAGCTGTGCTGTAAGTGAAGTGAGGCGGCAGGTGTTGAAAG
TCGATGTAGTTCGTAGGTCAGTTGATGTCGATGTGAAATGCTGATGCTAGT
GGACAGGGTGACTAGTGAATCGATGCTAGCCTAGCTAGTCAGTGGTGCTA
GCTACGATCGATTTCAGGCTGCTGTGGGTGCATAGCTGTGCTGTAAGTGAA
GTGAGGCGGCAGGTGTTGAAAGTCGATGTAGTTCGTAGGTCAGTTGATGTC
GATGTGAAATGCTGATGCTAGTGGACAGGGTGACTAGTGAATCGATGCTAG
CCTAGCTAGTCAGTGGTGCTAGCTACGATCGATTTCAGGCTGCTGTGGGTG
CATAGCTGTGCTGTAAGTGAAGTGAGGCGGCAGGTGTTGAAAGTCGATGTA
GTTCGTAGGTCAGTTGATGTCGATGTGAAATGCTGATGCTAGTGGACAGGG
TGACTAGTGAATCGATGCTAGCCTAGCTAGTCAGTGGTGCTAGCTACGATC
GATTTCAGGCTGCTGTGGGTGCATAGCTGTGCTGTAAGTGAAGTGAGGCGG
CAGGTGTTGAAAGTCGATGTAGTTCGTAGGTCAGTTGATGTCGATGTGAAAT
GCTGATGCTAGTGGACAGGGTGACTAGTGAATCGATGCTAGCCTAGCTAGT
CAGTGGTGCTAGCTACGATCGATTTCAGGCTGCT CCTAGCTAGTCAGTGGT
GTTCGTAGGTCAGTTGATGTCGATGTGAAATGCTGATGCTAGTGGACAGGG
TGACTAGTGAATCGATGCTAGCCTAGCTAGTCAGTGGTGCTAGCTACGATC
GATTTCAGGCTGCTGTGGGTGCATAGCTGTGCTGTAAGTGAAGTGAGGCGG
CAGGTGTTGAAAGTCGATGTAGTTCGTAGGTCAGTTGATGTCGATGTGA TC
GATGTAGTTCGTAGGTCAGTTGATGTCGATGTGAAATGCTGATGCTAGTGGA
CAGGGTGACTAGTGAATCGATGCTAGCCTAGCTAGTCAGTGGTGCTATTGT
GCTACGATCGATTTCAGGCTGCTGTGGGTGCATAGCTGTGCTGTAAGTGAA
GTGAGGCGGCAGGTGTTGAAAGTCGATGTAGTTCGTAGGTCAGTTGATGTC
GATGTGAAATGCTGATGCTAGTGGACAGGGTGACTAGTGAATCGATGCTAG
CCTAGCTAGTCAGTGGTGCTAGCTACGATCGATTTCAGGCTGCTGTGGGTG

GATGTGAAATGCTGATGCTAGTGGACAGGGTGACTAGTGAATCGATGCTAG
CCTAGCTAGTCAGTGGTGCTAGCTACGATCGATTTCAGGCTGCTGTGGGTG
CATAGCTGTGCTGTAAGTGAAGTGAGGCGGCAGGTGTTGAAAGTCGATGTA
GTTCGTAGGTCAGTTGATGTCGATGTGAAATGCTGATGCTAGTGGACAGGG
TGACTAGTGAATCGATGCTAGCCTAGCTAGTCAGTGGTGCTAGCTACGATC
GATTTCAGGCTGCTGTGGGTGCATAGCTGTGCTGTAAGTGAAGTGAGGCGG
CAGGTGTTGAAAGTCGATGTAGTTCGTAGGTCAGTTGATGTCGATGTGAAAT
GCTGATGCTAGTGGACAGGGTGACTAGTGAATCGATGCTAGCCTAGCTAGT
CAGTGGTGCTAGCTACGATCGATTTCAGGCTGCT CCTAGCTAGTCAGTGGT
GTTCGTAGGTCAGTTGATGTCGATGTGAAATGCTGATGCTAGTGGACAGGG
TGACTAGTGAATCGATGCTAGCCTAGCTAGTCAGTGGTGCTAGCTACGATC
GATTTCAGGCTGCTGTGGGTGCATAGCTGTGCTGTAAGTGAAGTGAGGCGG
CAGGTGTTGAAAGTCGATGTAGTTCGTAGGTCAGTTGATGTCGATGTGA TC
GATGTAGTTCGTAGGTCAGTTGATGTCGATGTGAAATGCTGATGCTAGTGGA
CAGGGTGACTAGTGAATCGATGCTAGCCTAGCTAGTCAGTGGTGCTATTGT
GCTACGATCGATTTCAGGCTGCTGTGGGTGCATAGCTGTGCTGTAAGTGAA
GTGAGGCGGCAGGTGTTGAAAGTCGATGTAGTTCGTAGGTCAGTTGATGTC
GATGTGAAATGCTGATGCTAGTGGACAGGGTGACTAGTGAATCGATGCTAG
CCTAGCTAGTCAGTGGTGCTAGCTACGATCGATTTCAGGCTGCTGTGGGTG
AAACGATGTGAAATGCTGATGCTAGTGGACAGGGTGACTAGTGAATCGATG
CTAGCCTAGCTAGTCAGTGGTGCTAGCTACGATCGATTTCAGGCTGCTGTG
GGTGGCTGATGCTAGTGGACAGGGTGACTAGTGAATCGATGCTAGCCTAGCT
AGTCAGTGGTGCTAGCTACGATCGATTTCAGGCTGCCCTAGCTAGTCAGTGG

CAGGTGTTGAAAGTCGATGTAGTTCGTAGGTCAGTTGATGTCGATGTGATC
GATGTAGTTCGTAGGTCAGTTGATGTCGATGTGAAATGCTGATGCTAGTGGA
CAGGGTGACTAGTGAATCGATGCTAGCCTAGCTAGTCAGTGGTGCTATTGT
GCTACGATCGATTTCAGGCTGCTGTGGGTGCATAGCTGTGCTGTAAGTGAA
GTGAGGCGGCAGGTGTTGAAAGTCGATGTAGTTCGTAGGTCAGTTGATGTC
GATGTGAAATGCTGATGCTAGTGGACAGGGTGACTAGTGAATCGATGCTAG
CCTAGCTAGTCAGTGGTGCTAGCTACGATCGATTTCAGGCTGCTGTGGGTG
AAACGATGTGAAATGCTGATGCTAGTGGACAGGGTGACTAGTGAATCGATG
CTAGCCTAGCTAGTCAGTGGTGCTAGCTACGATCGATTTCAGGCTGCTGTG
GCTGATGCTAGTGGACAGGGTGACTAGTGAATCGATGCTAGCCTAGCTAGT
CAGTGGTGCTAGCTACGATCGATTTCAGGCTGCT CCTAGCTAGTCAGTGGT
GATGTGAAATGCTGATGCTAGTGGACAGGGTGACTAGTGAATCGATGCTAG
CCTAGCTAGTCAGTGGTGCTAGCTACGATCGATTTCAGGCTGCTGTGGGTG
CATAGCTGTGCTGTAAGTGAAGTGAGGCGGCAGGTGTTGAAAGTCGATGTA
GTTCGTAGGTCAGTTGATGTCGATGTGAAATGCTGATGCTAGTGGACAGGG
TGACTAGTGAATCGATGCTAGCCTAGCTAGTCAGTGGTGCTAGCTACGATC
GATTTCAGGCTGCTGTGGGTGCATAGCTGTGCTGTAAGTGAAGTGAGGCGG
CAGGTGTTGAAAGTCGATGTAGTTCGTAGGTCAGTTGATGTCGATGTGAAAT
GCTGATGCTAGTGGACAGGGTGACTAGTGAATCGATGCTAGCCTAGCTAGT
CAGTGGTGCTAGCTACGATCGATTTCAGGCTGCT CCTAGCTAGTCAGTGGT
GTTCGTAGGTCAGTTGATGTCGATGTGAAATGCTGATGCTAGTGGACAGGG
TGACTAGTGAATCGATGCTAGCCTAGCTAGTCAGTGGTGCTAGCTACGATC
GATTTCAGGCTGCTGTGGGTGCATAGCTGTGCTGTAAGTGAAGTGAGGCGG

… and ~ 3,000,000 more pages!

(one month to show 24/7)

Just four letter:

~715 Mb

DNA code is billion times more efficient

A

C

G

T

2 bits

0 0

1 1

8 bits = 1b

4/8*3*10^9

Differences in the code are important

Among unrelated individuals, 99.4% of the sequence is similar
That is still over 1,000,000 differences.

You and chimpanzee: 99%

Advanced diagnostics
(early detection and, possibly, prevention of 4,000 genetic disorders)

Research instrumentation
(reconstruction of the tree of life, human history, psychology)

Cancer: disease of DNA

Personal pharmaceutics
(tailor drugs to an individual's genetic make-up)

Prenatal diagnostics

Single cell sequencing

The Sanger's method

Nobel Prize in Chemistry 1980

As the DNA is synthesized, nucleotides are added on to the growing chain by the DNA polymerase.

The reactions start from the same nucleotide and end with a specific base

dye label

chain termination with ddGTP

5'-GAATGTCCTTTCTCTAAGTCCTAAGTCCTCCG
3'-GGAGACTTACAGGAAAGAGATTCAGGATTCAGGAGGCCTACCATGAAGATCAAG-5'

chain termination with ddATP

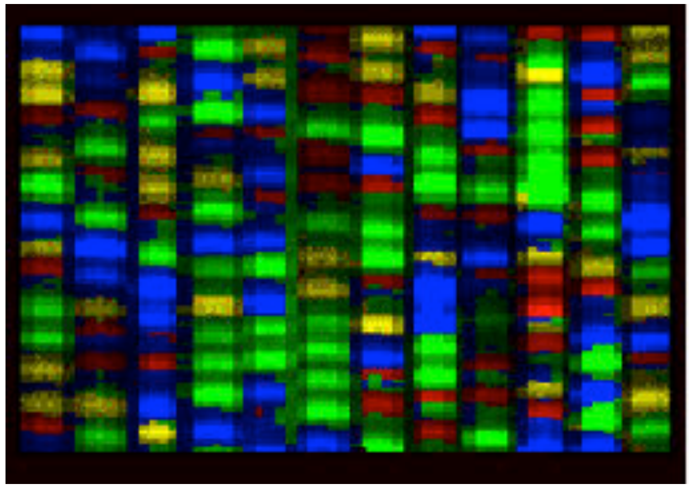
5'-GAATGTCCTTTCTCTAAGTCCTAAGTCCTCCGGA
3'-GGAGACTTACAGGAAAGAGATTCAGGATTCAGGAGGCCTACCATGAAGATCAAG-5'

chain termination with ddTTP

5'-GAATGTCCTTTCTCTAAGTCCTAAGTCCT
3'-GGAGACTTACAGGAAAGAGATTCAGGATTCAGGAGGCCTACCATGAAGATCAAG-5'

chain termination with ddCTP

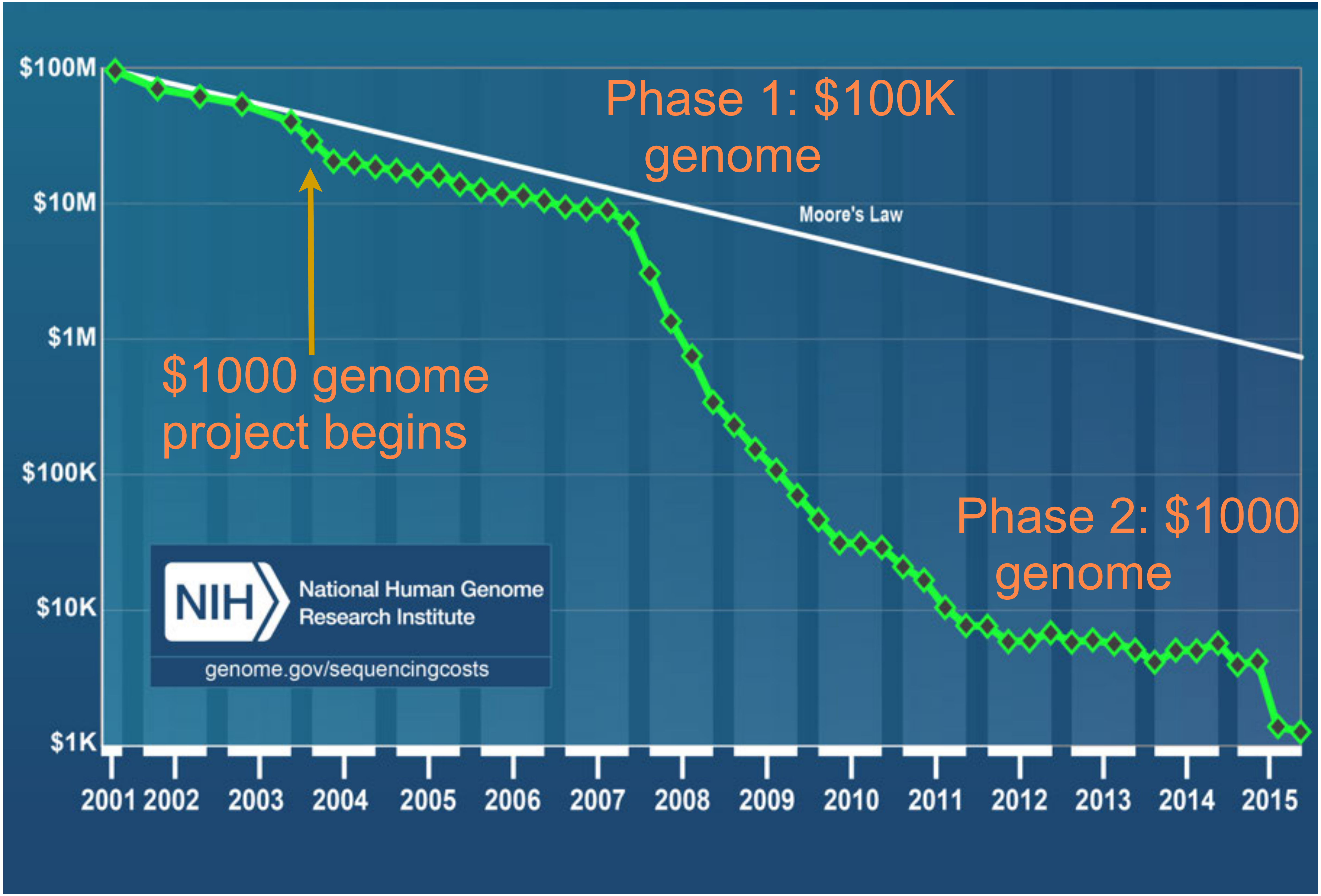
5'-GAATGTCCTTTCTCTAAGTCCTAAGTCC
3'-GGAGACTTACAGGAAAGAGATTCAGGATTCAGGAGGCCTACCATGAAGATCAAG-5'



Fluorescence-based sequence gel

http://bbrp.llnl.gov

Cost of sequencing a human genome
(logarithmic scale)

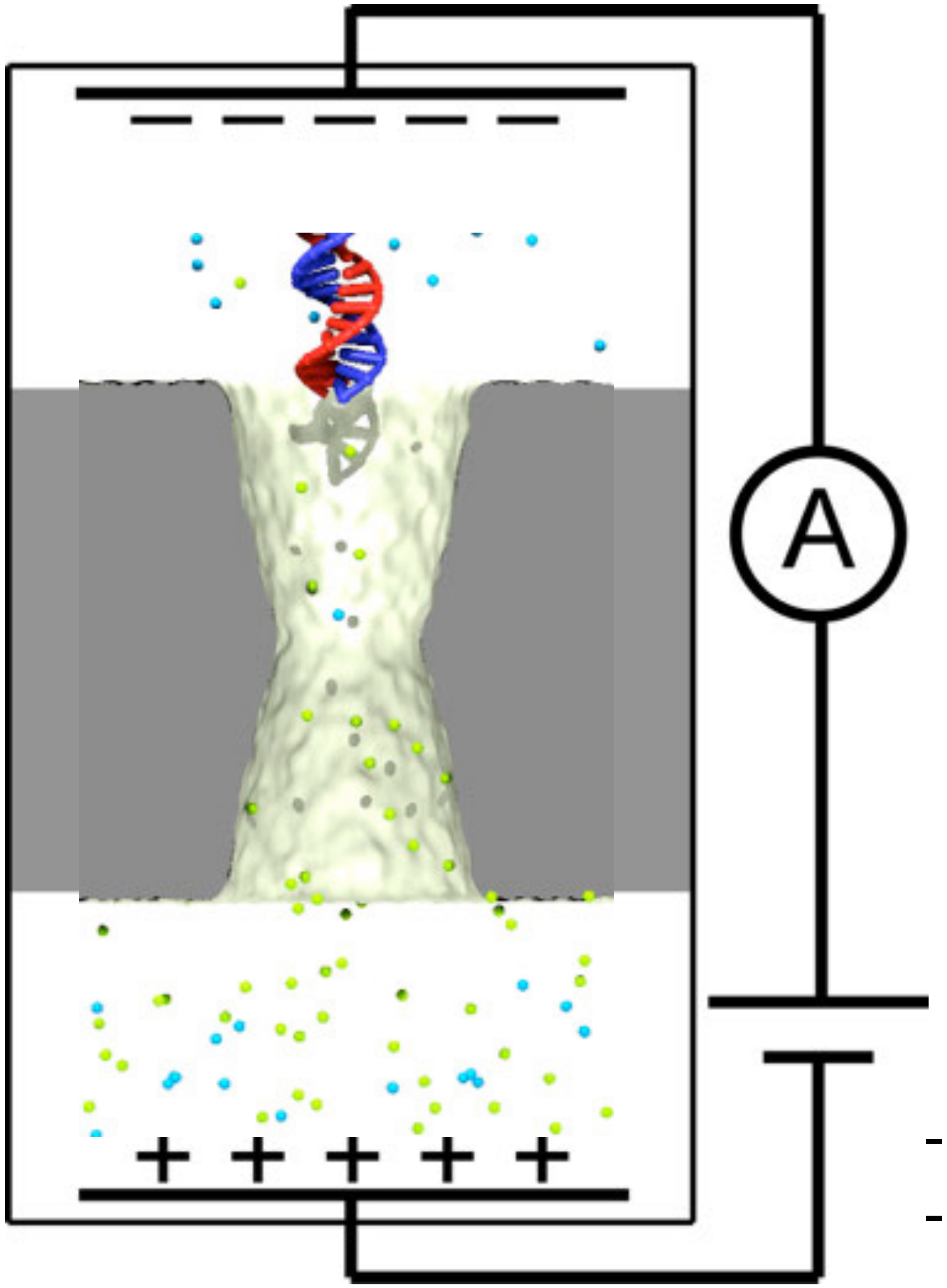


Phase 1: $100K genome

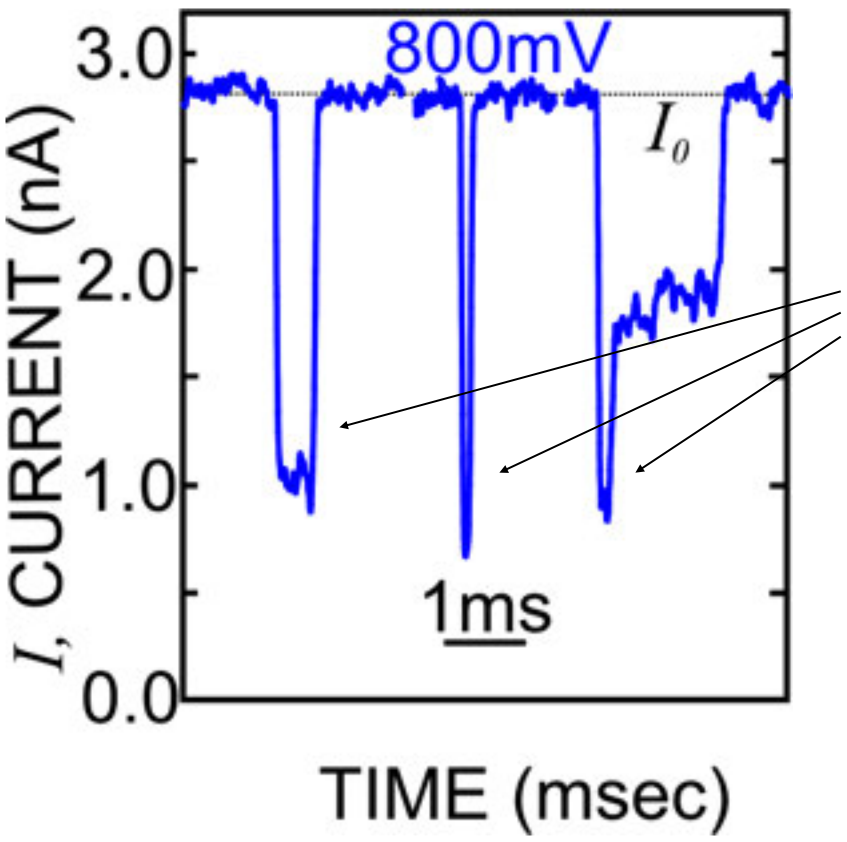
$1000 genome project begins

Phase 2: $1000 genome

DNA electrophoresis through a nanopore



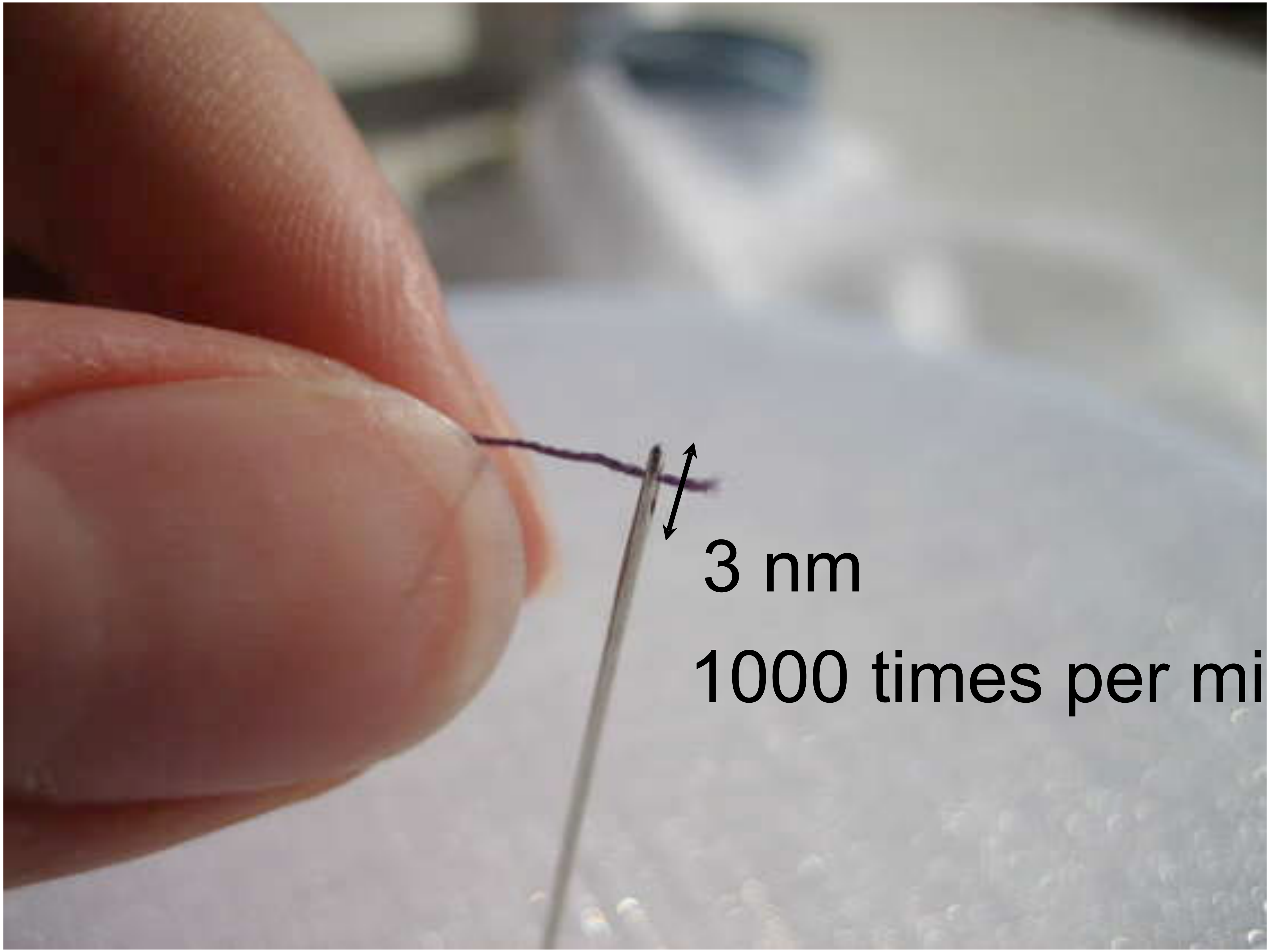
Ionic current through pore measured



current transients associated with passage of dsDNA

- Isolates 1nm^3 of volume
- Automatic loading and reloading
- Highly processive, single-file transport
- Compatible with several detection schemes
- No limit on the read length

The nanopore technology

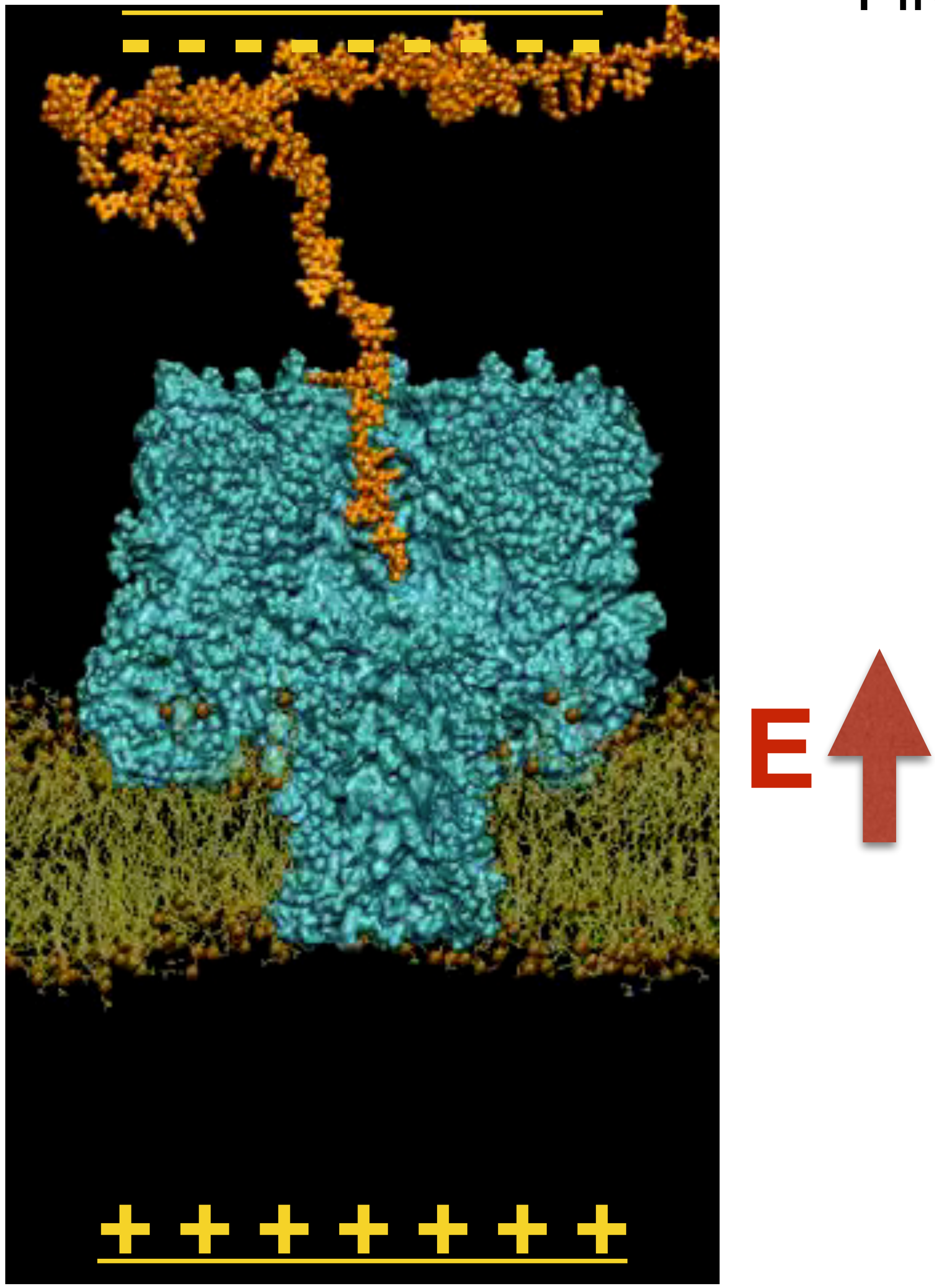


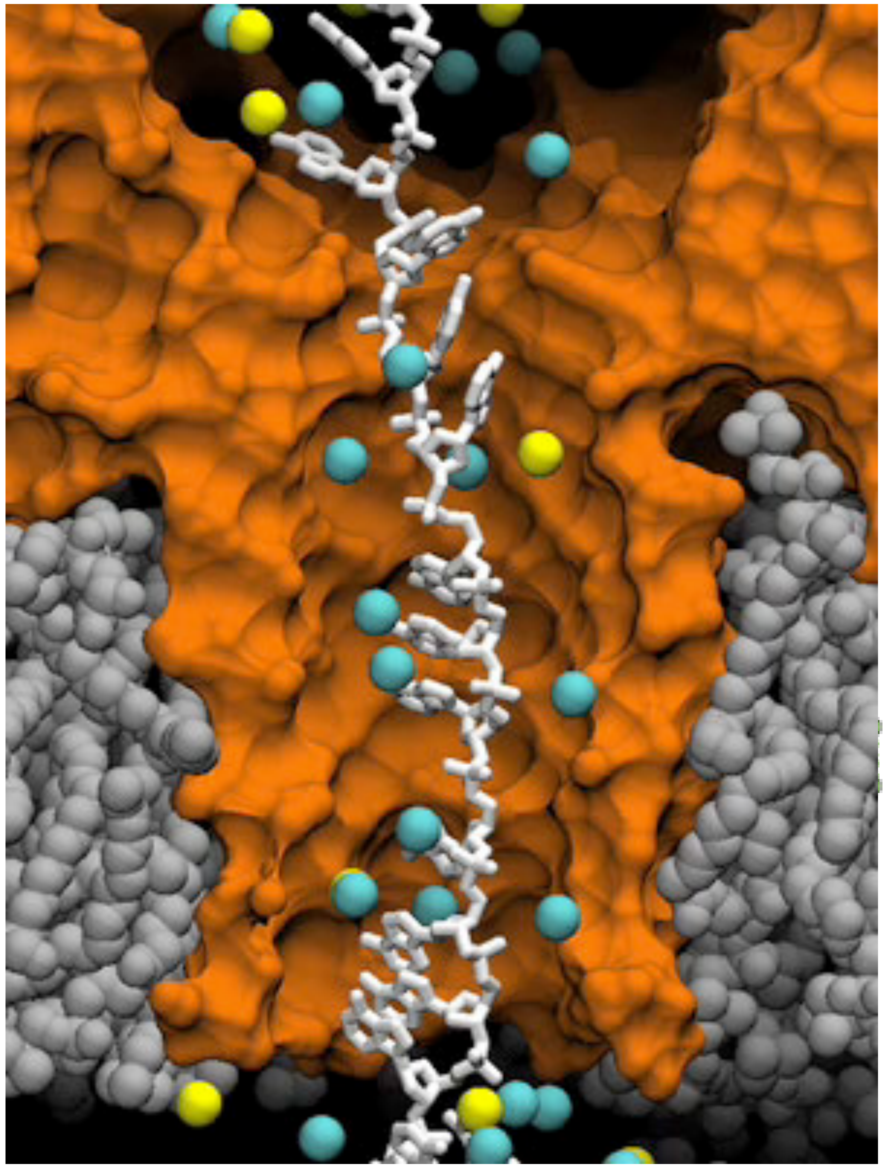
3 nm

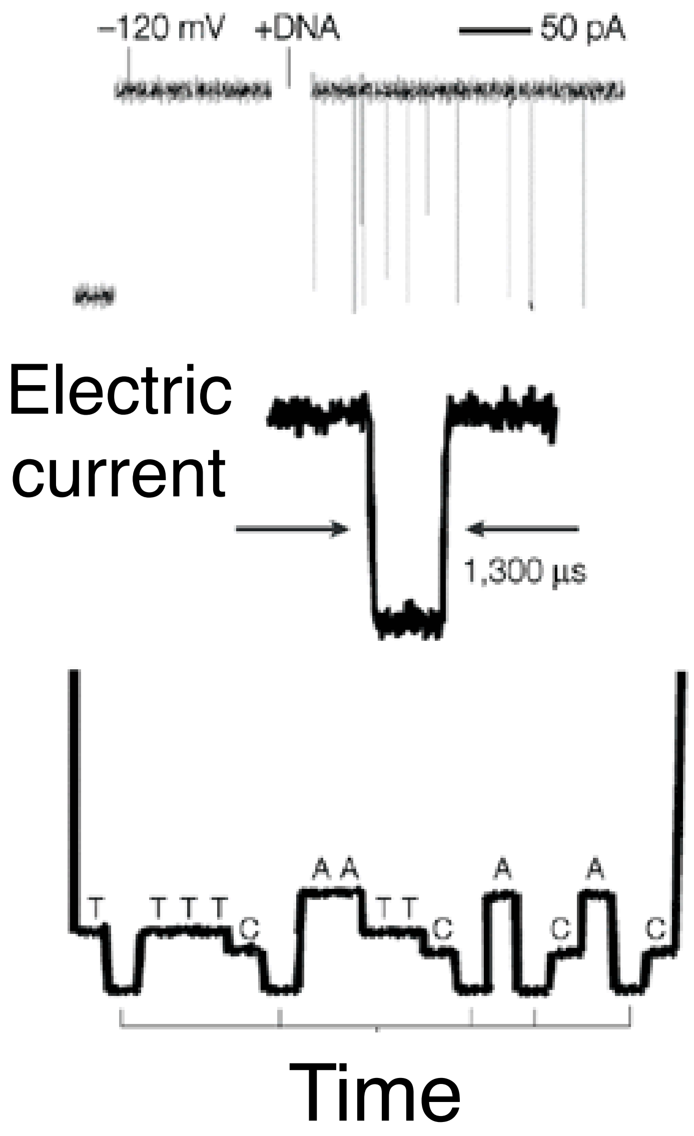
1000 times per minute

Nanopore sequencing of DNA

First described by Kasianowicz et al. PNAS 1996



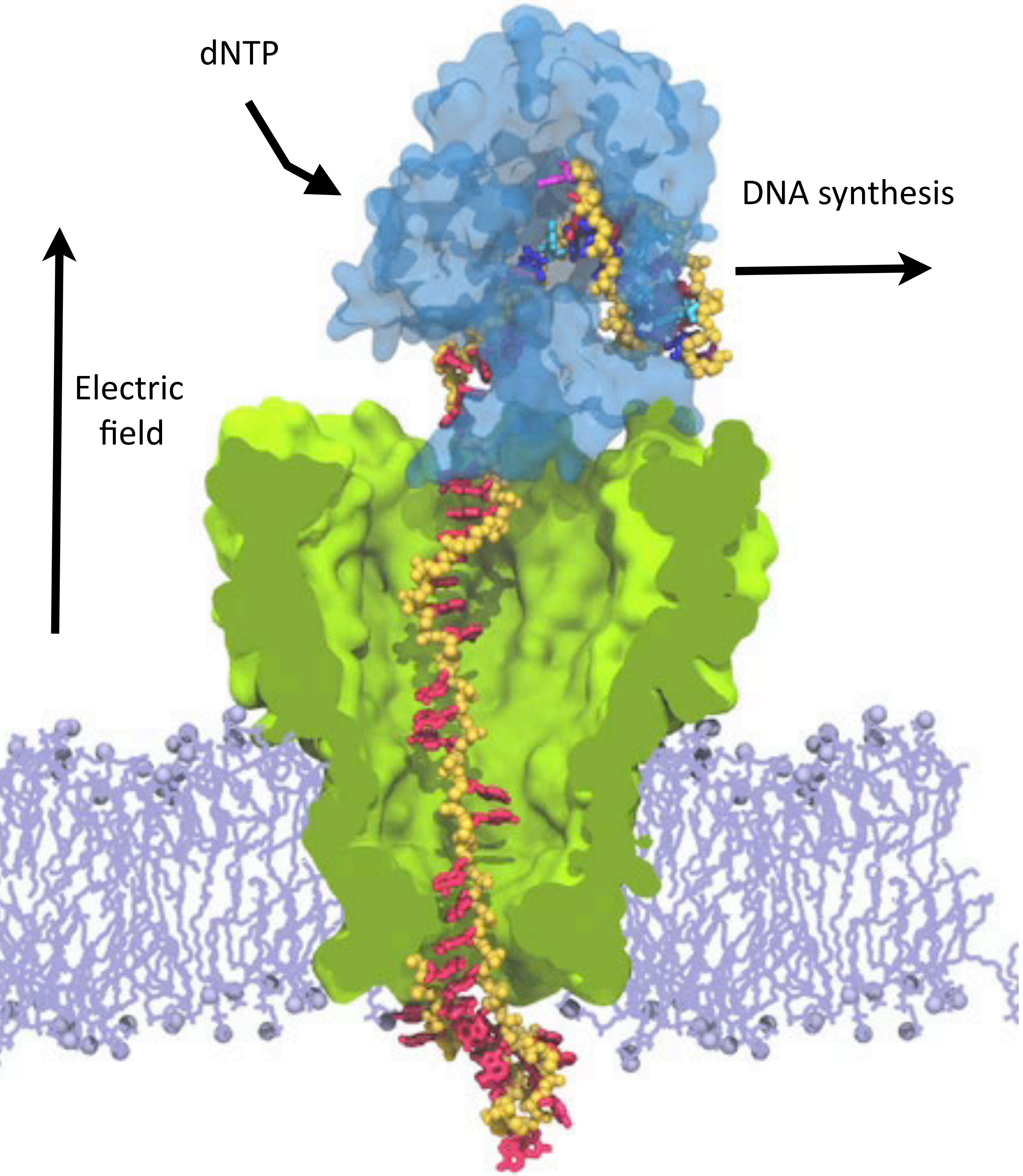




Nature Reviews Drug Discovery 1, 77-84 (January 2002)

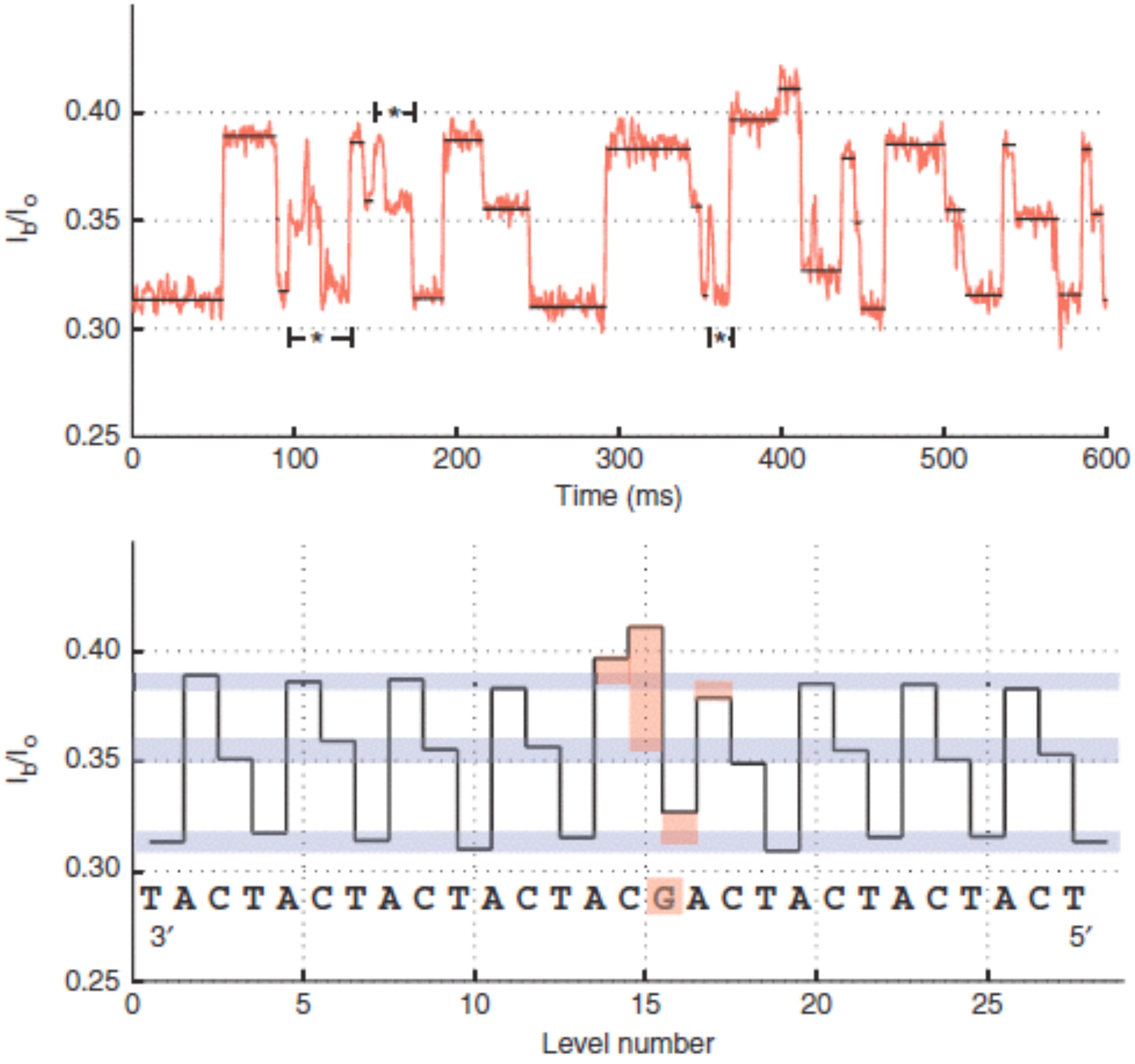
The ionic current blockade reveals the sequence of the confined nucleotides

Sequencing DNA using a biological nanopore



MD simulation ssDNA- DNA polymeraze complex

Experimentally measured ionic current blockades



Nature Biotech. 30: 349 - 353 (2012)

Nature Biotech. 32: 829 - 834 (2014)

MinION (Oxford Nanopore Technologies)

MinION: 800 parallel detection wells

Read length: up to 200,000 nucleotides

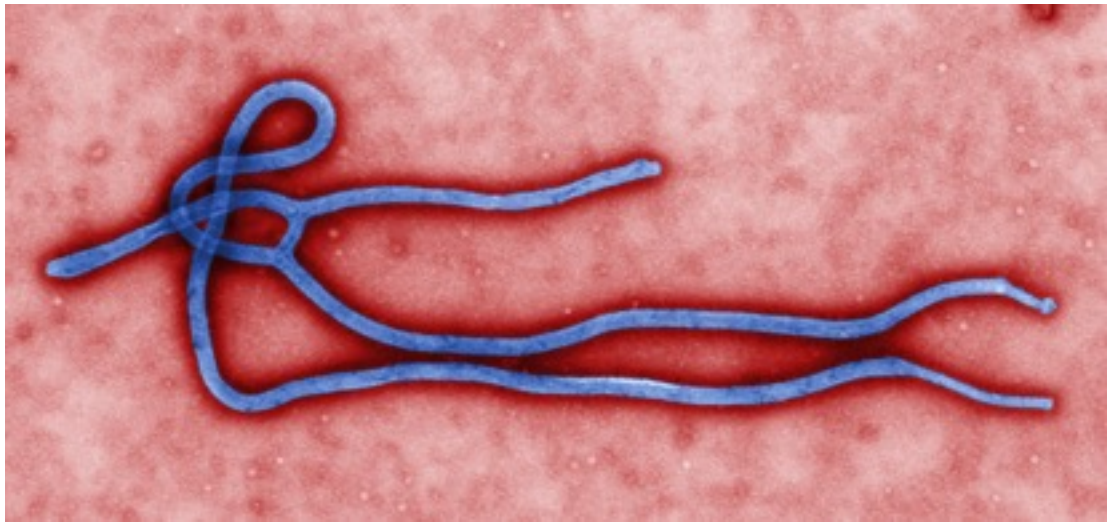
Biological pore (R9, CsgG-derivative)

Helicase motor to control translocation

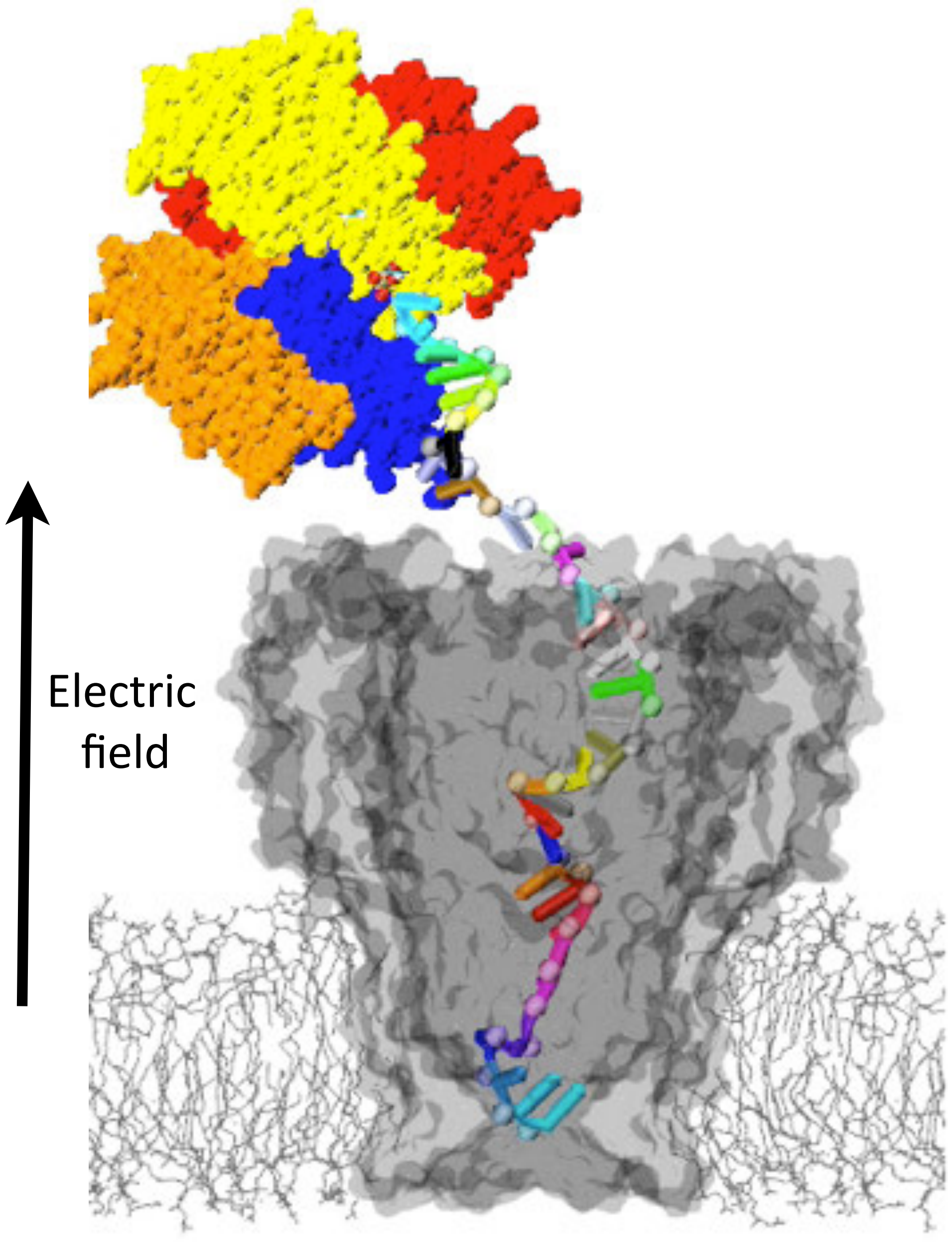
Direct readout of DNA epigenetic markers

Accuracy: 80 - 98 %

Only 100g, great for *in situ* measurements

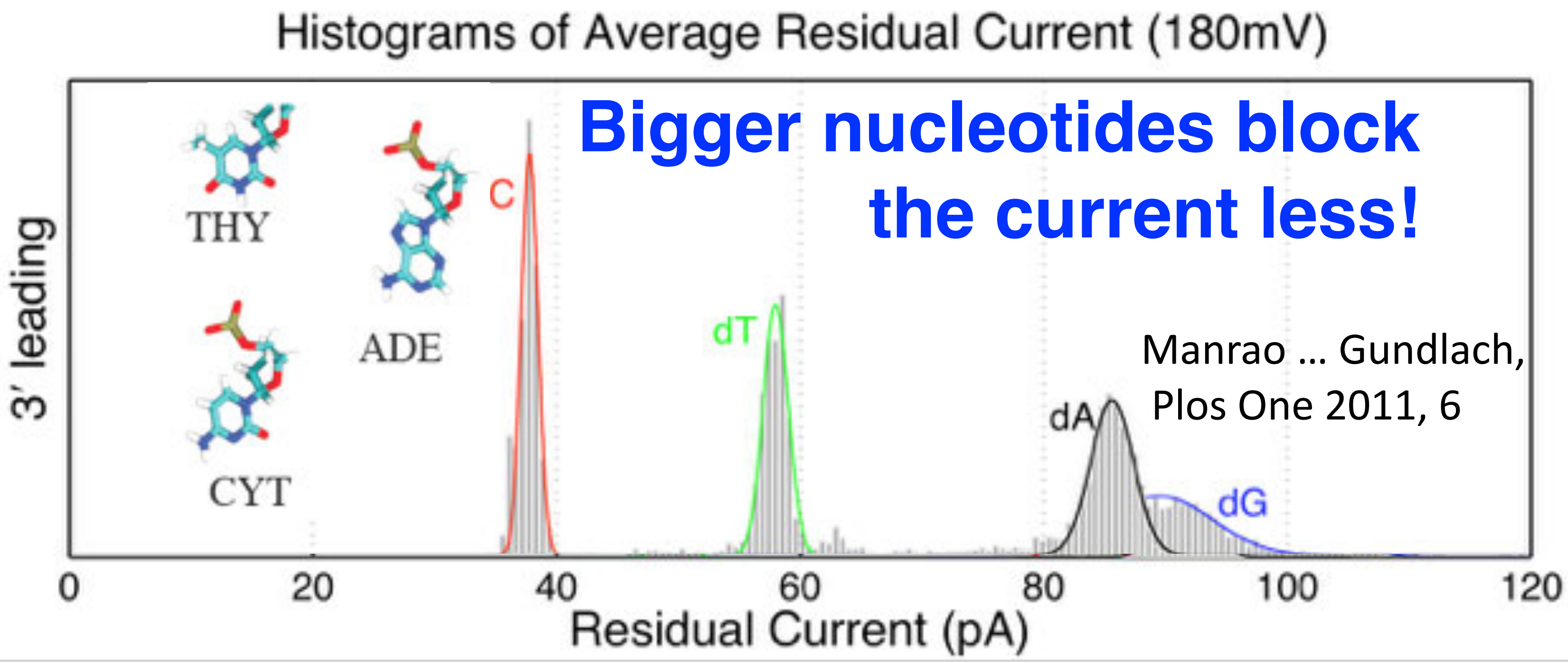


Biological nanopore: why does it work?



Electric field

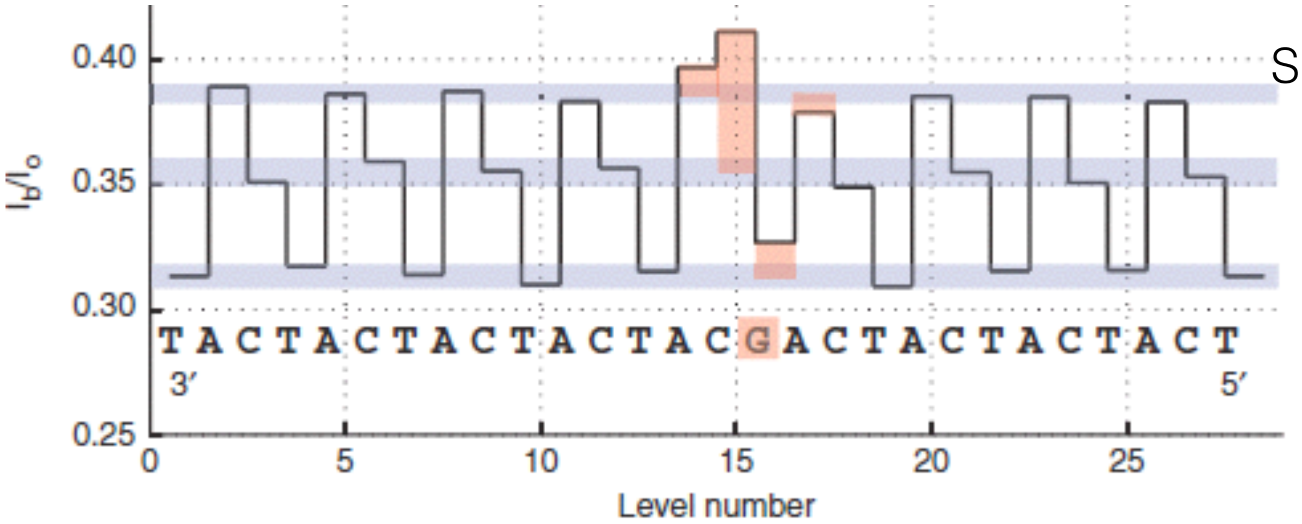
MD simulation neutravidin-anchored ssDNA in MspA



Bigger nucleotides block the current less!

Manrao … Gundlach, Plos One 2011, 6

Blockades are non-additive



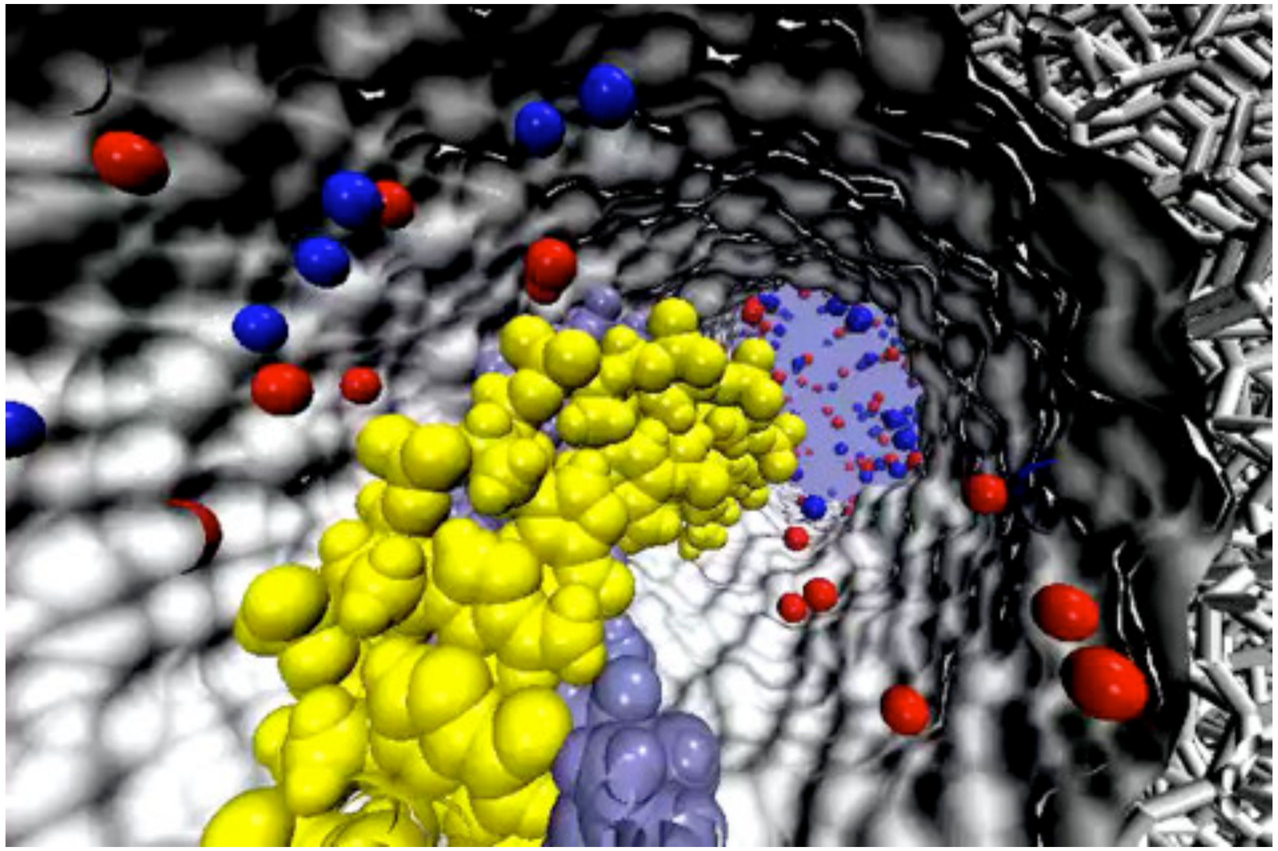
Single G substitution can both increase and decrees the current !

Manrao … Gundlach, Nature Biotech. 30: 349 - 353 (2012)

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

<u>Time scale</u>: ~ 0.1-100 μs

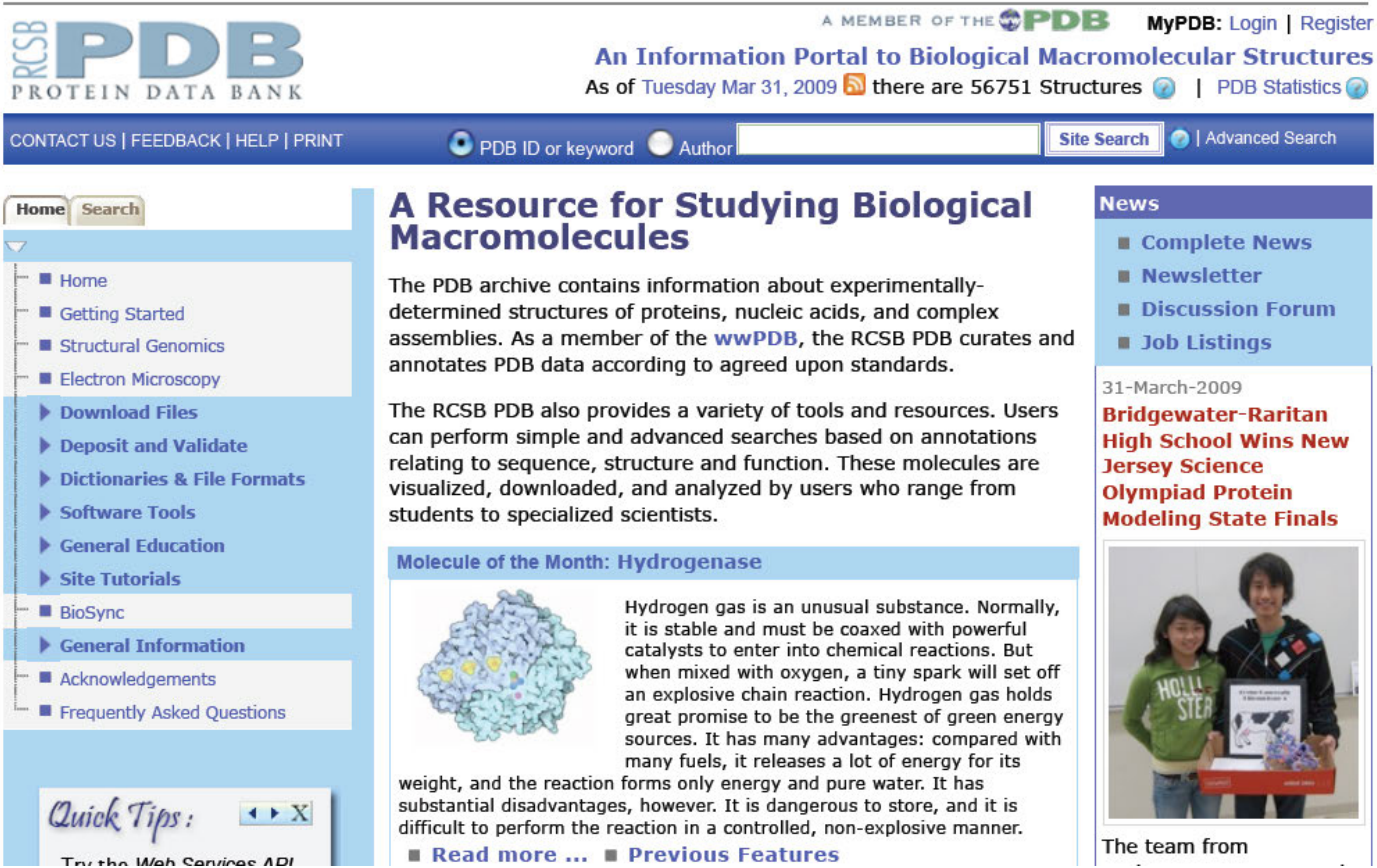
<u>Length scale</u>: 10K - 100M atoms or (< 50 nm)^3

<u>Time resolution</u>: 2 fs

<u>Spacial resolution:</u> 0.1 A

Nanoscale 2:468 (2010)

Setting up a simulation is like cooking



An Information Portal to Biological Macromolecular Structures

As of Tuesday Mar 31, 2009 there are 56751 Structures | PDB Statistics

A Resource for Studying Biological Macromolecules

The PDB archive contains information about experimentally-determined structures of proteins, nucleic acids, and complex assemblies. As a member of the **wwPDB**, the RCSB PDB curates and annotates PDB data according to agreed upon standards.

The RCSB PDB also provides a variety of tools and resources. Users can perform simple and advanced searches based on annotations relating to sequence, structure and function. These molecules are visualized, downloaded, and analyzed by users who range from students to specialized scientists.

Molecule of the Month: Hydrogenase

Hydrogen gas is an unusual substance. Normally, it is stable and must be coaxed with powerful catalysts to enter into chemical reactions. But when mixed with oxygen, a tiny spark will set off an explosive chain reaction. Hydrogen gas holds great promise to be the greenest of green energy sources. It has many advantages: compared with many fuels, it releases a lot of energy for its weight, and the reaction forms only energy and pure water. It has substantial disadvantages, however. It is dangerous to store, and it is difficult to perform the reaction in a controlled, non-explosive manner.

31-March-2009

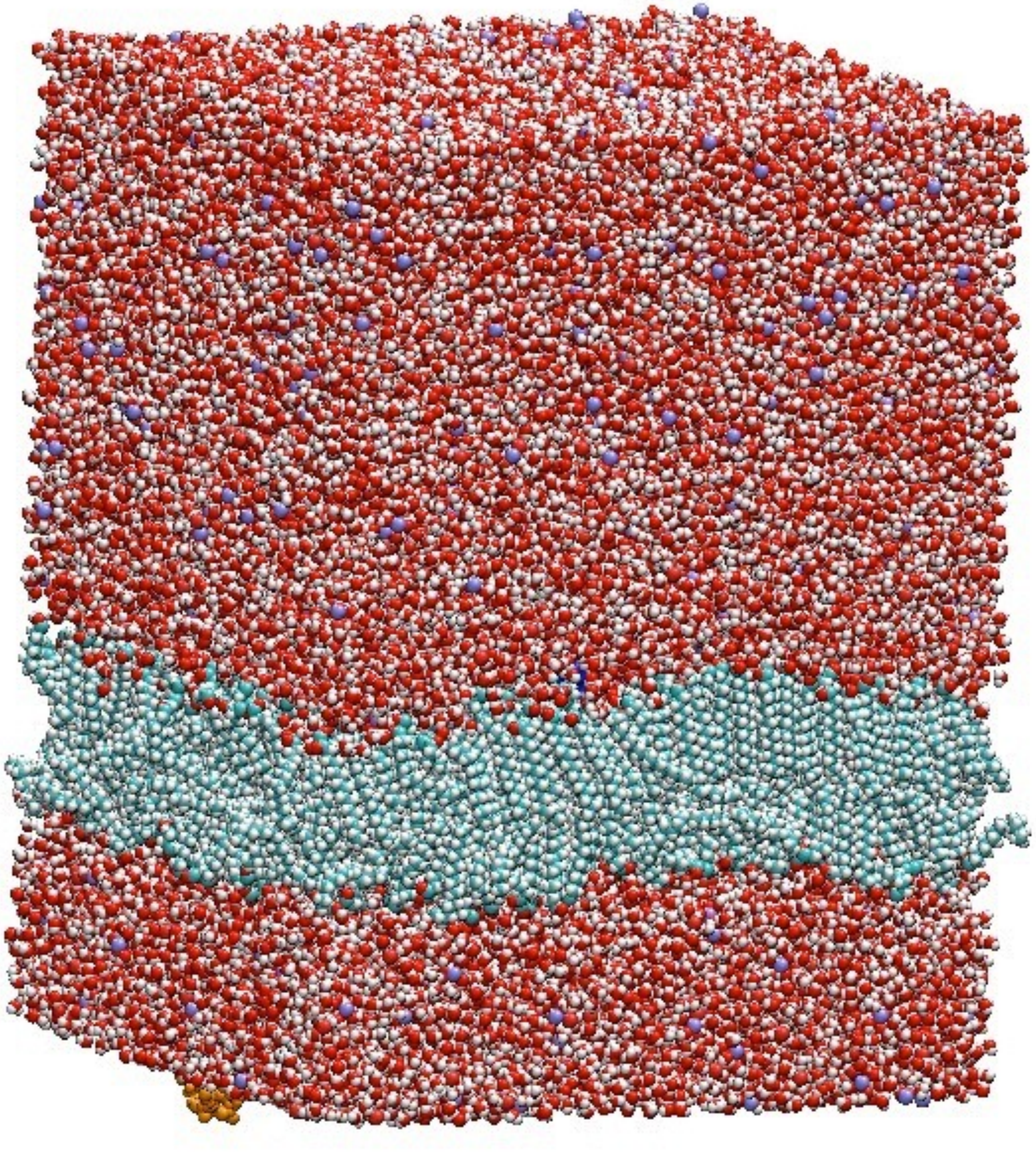
Bridgewater-Raritan High School Wins New Jersey Science Olympiad Protein Modeling State Finals

The team from

Quick Tips:

Try the Web Services API

Setting up a simulation is like cooking



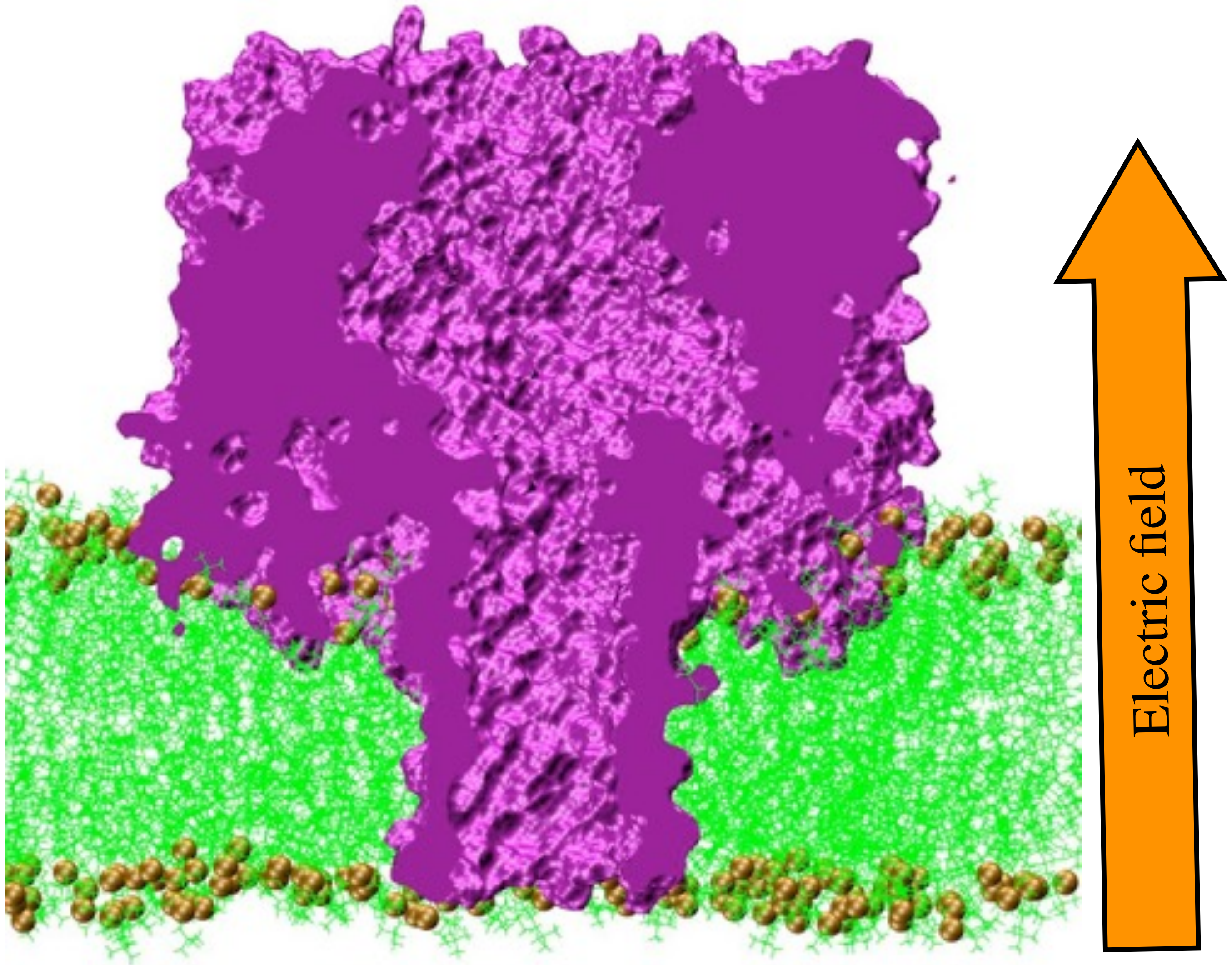
Components

- protein
- DNA
- lipid
- ions
- water

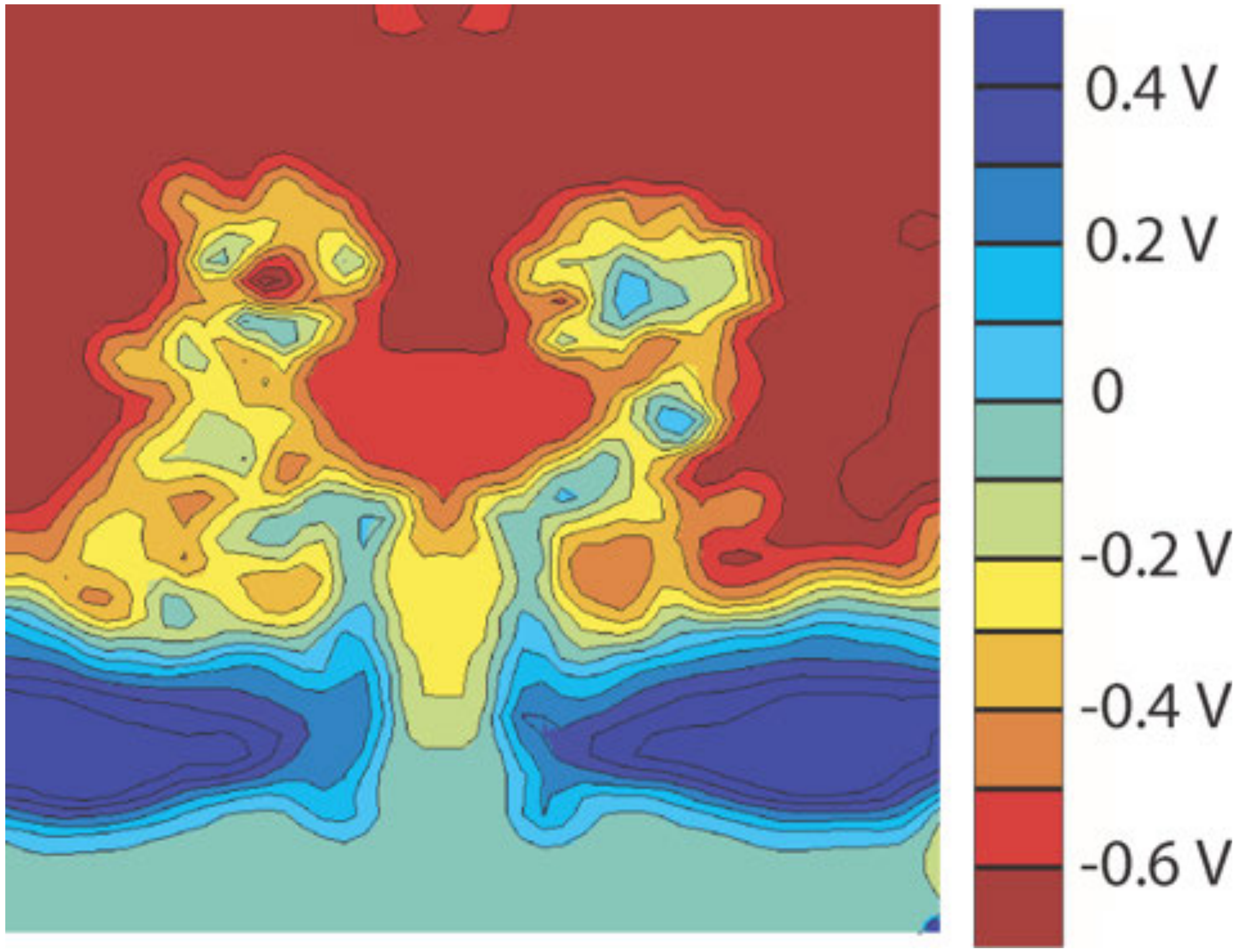
F = ma @ 300 K

Time step = 1 fs

Computing conductance of α-hemolysin with molecular dynamics



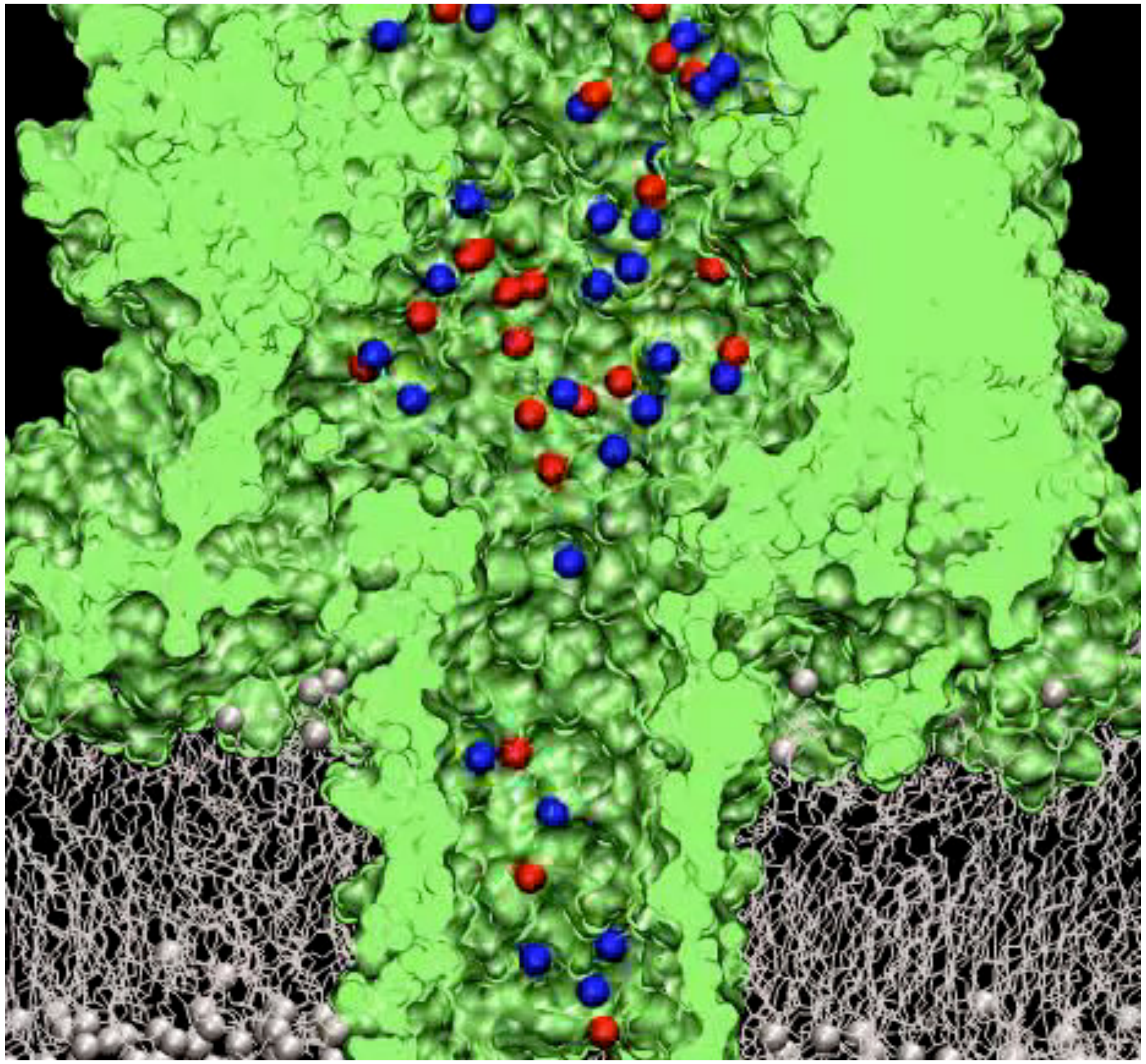
Protein + lipid bilayer membrane + 1M water solution of KCl = ~300,000 atoms

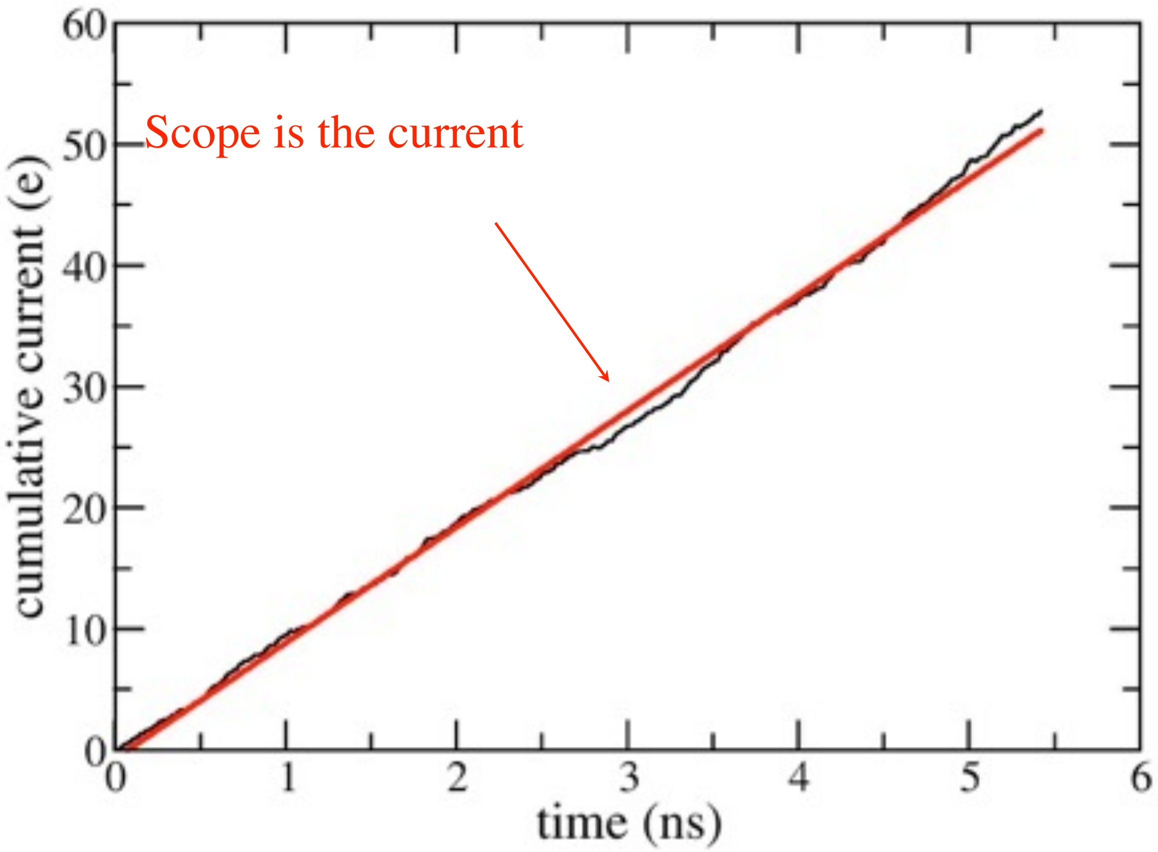


Average electrostatic potential map

Current-voltage curve of α-hemolysin

Biophys. J. 88:3745 (2005)





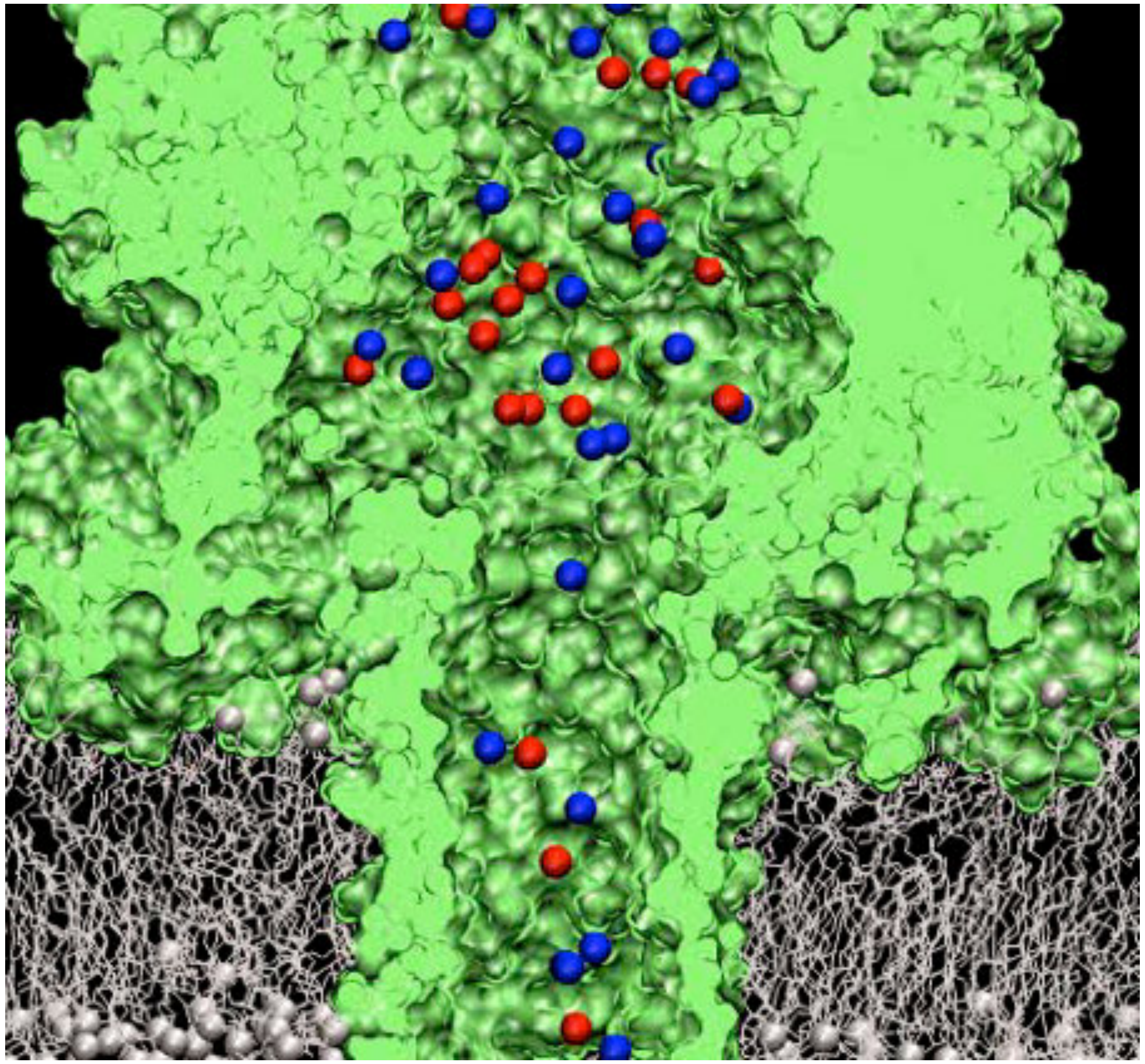
Scope is the current

$$I(t) = \frac{1}{\Delta t L_z} \sum_{i=1}^{N} q_i (z_i(t + \Delta t) - z_i(t))$$

Instantaneous current

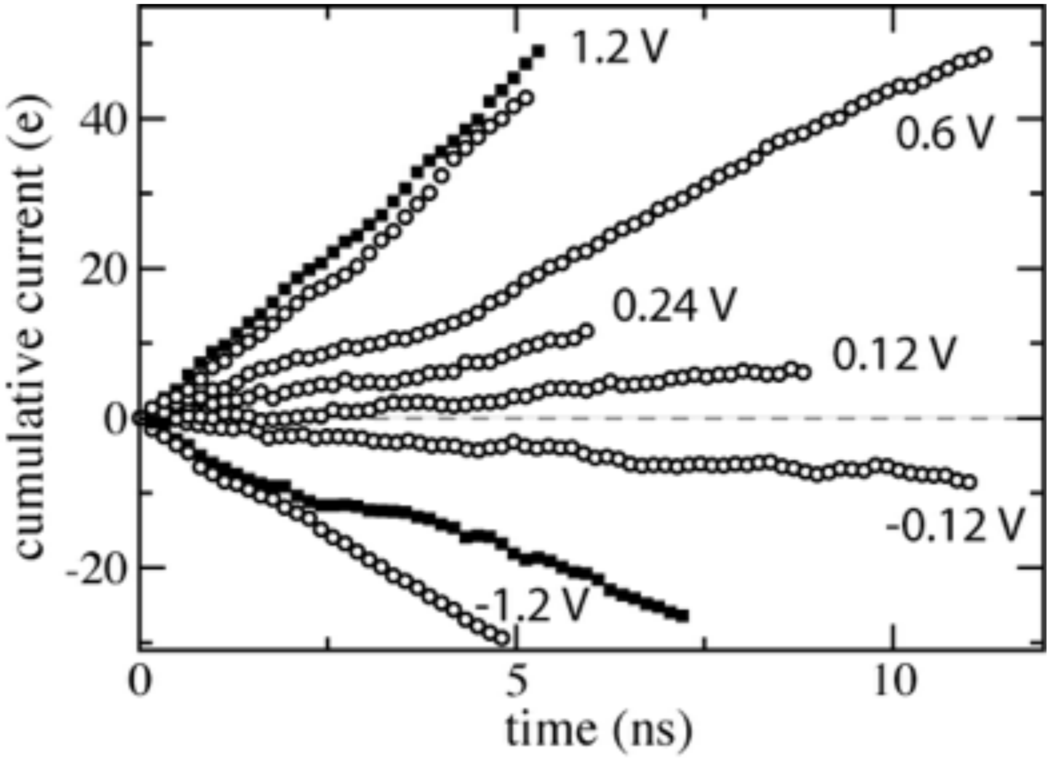
Current-voltage curve of α-hemolysin

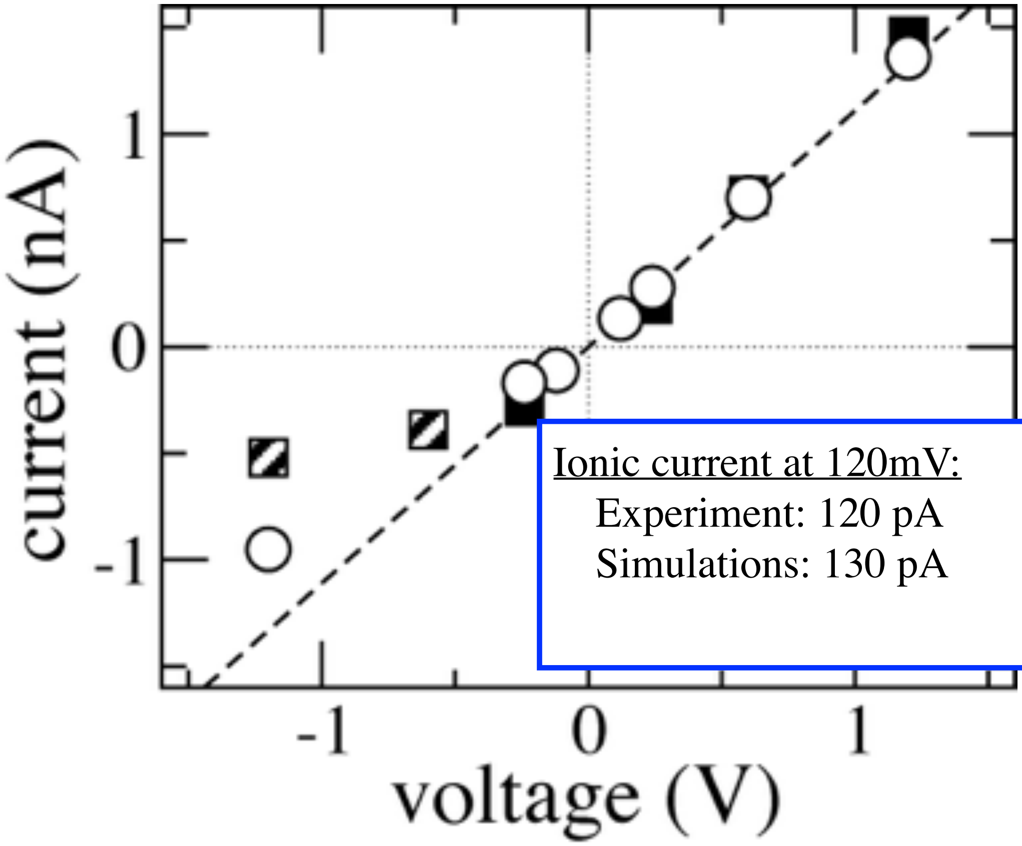
Biophys. J. 88:3745 (2005)



$$I(t) = \frac{1}{\Delta t L_z} \sum_{i=1}^{N} q_i (z_i(t + \Delta t) - z_i(t))$$

Instantaneous current

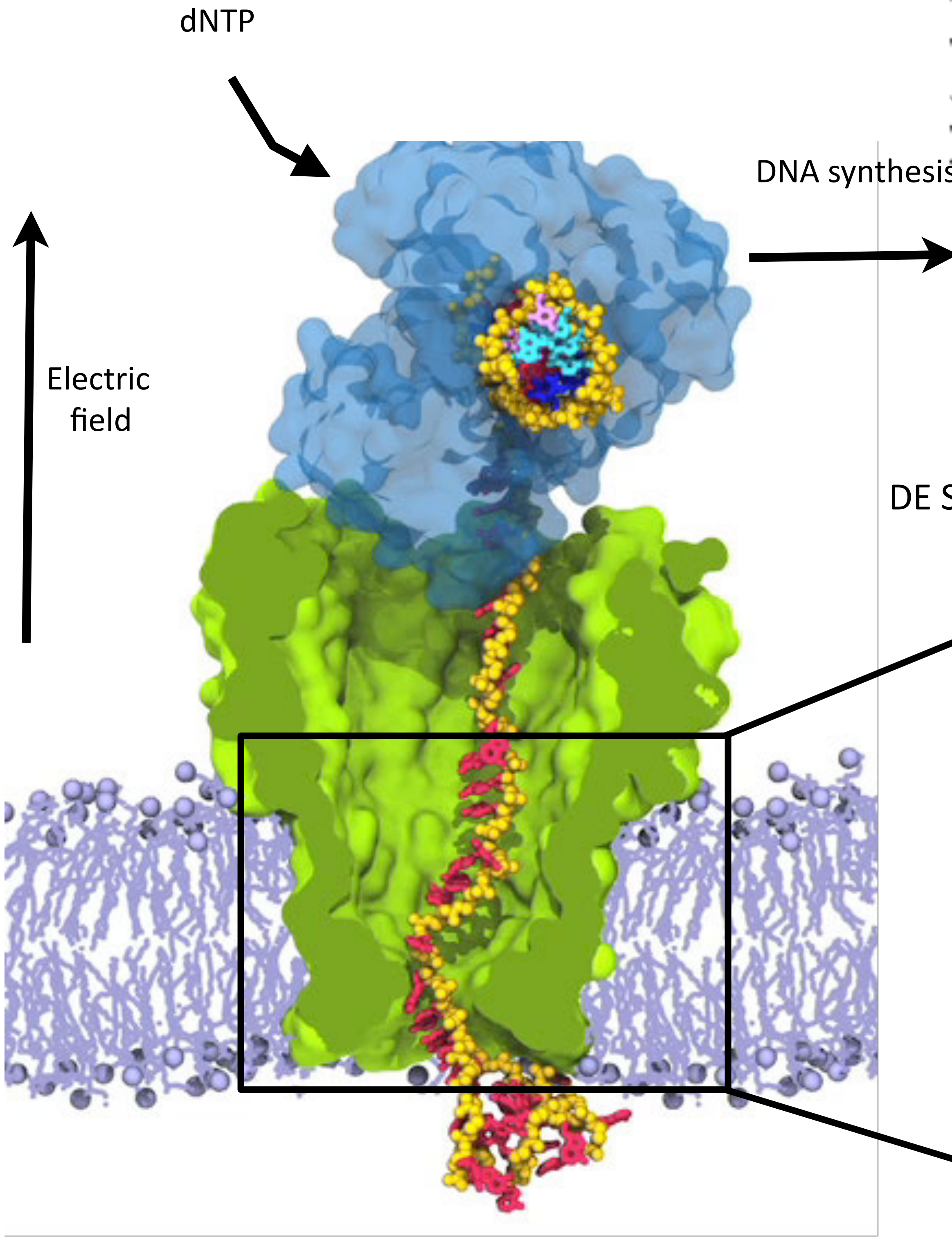


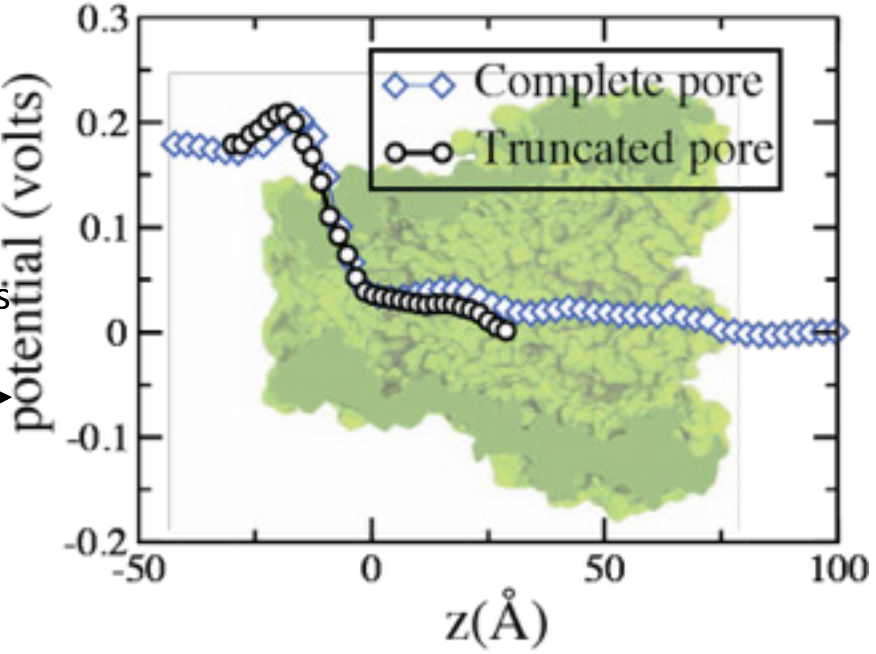


Ionic current at 120mV:
Experiment: 120 pA
Simulations: 130 pA

MD simulations of current blockades in MspA

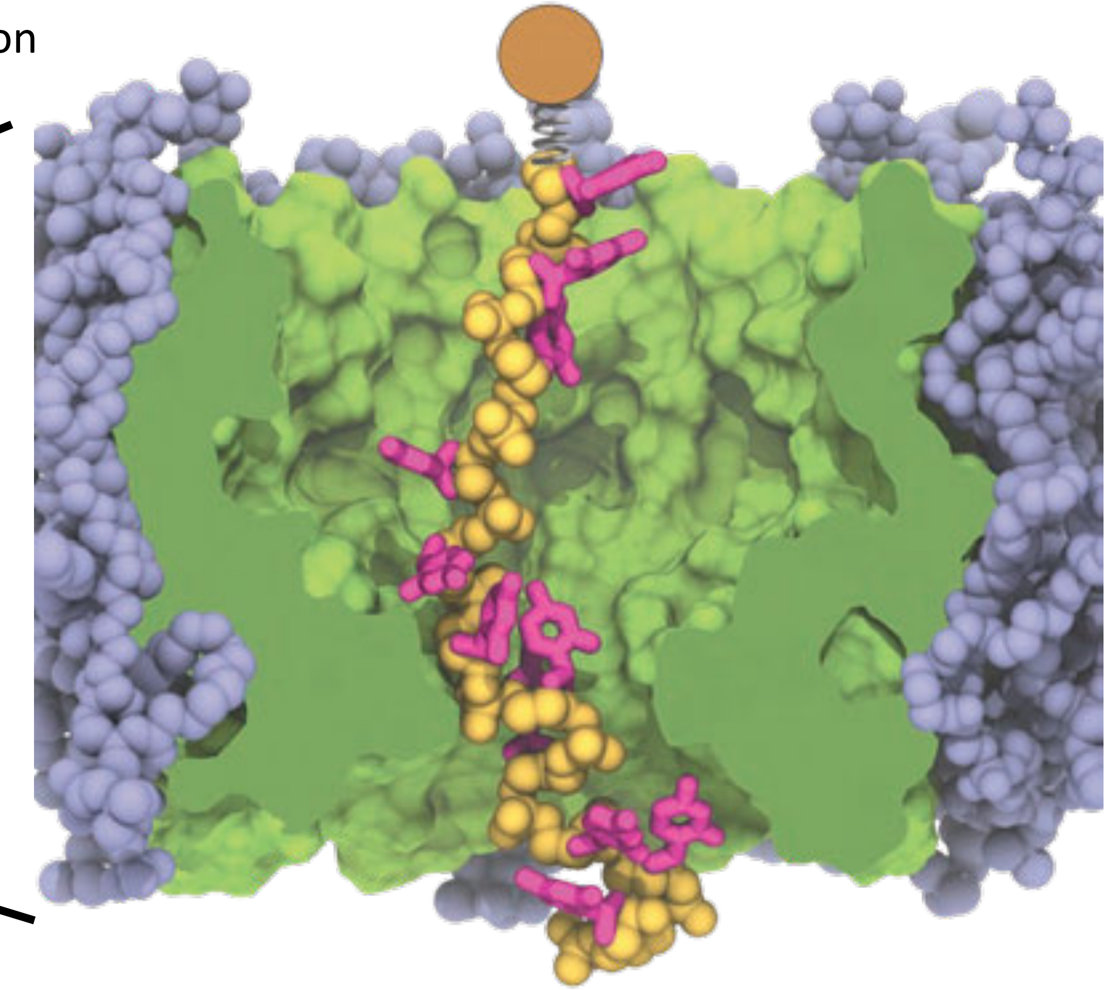
ACS Nano 2016, 10, 4644





Swati Bhattacharya

DE Shaw's Anton



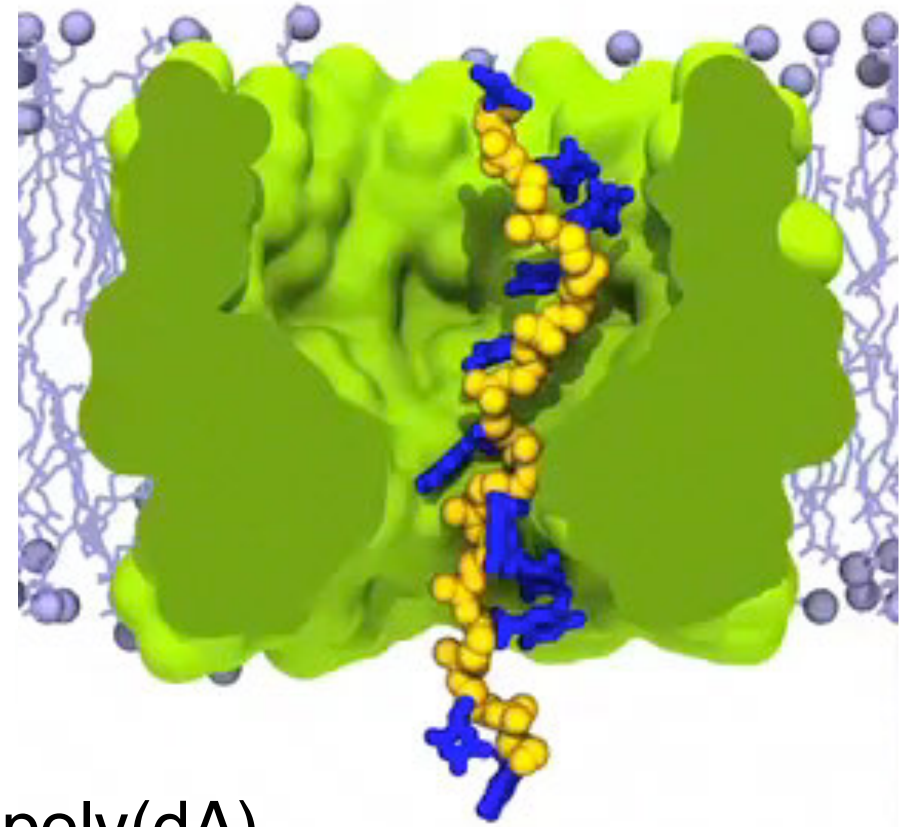
MD simulation ssDNA- DNA polymerase complex
(350,000 atoms, 150 ns)

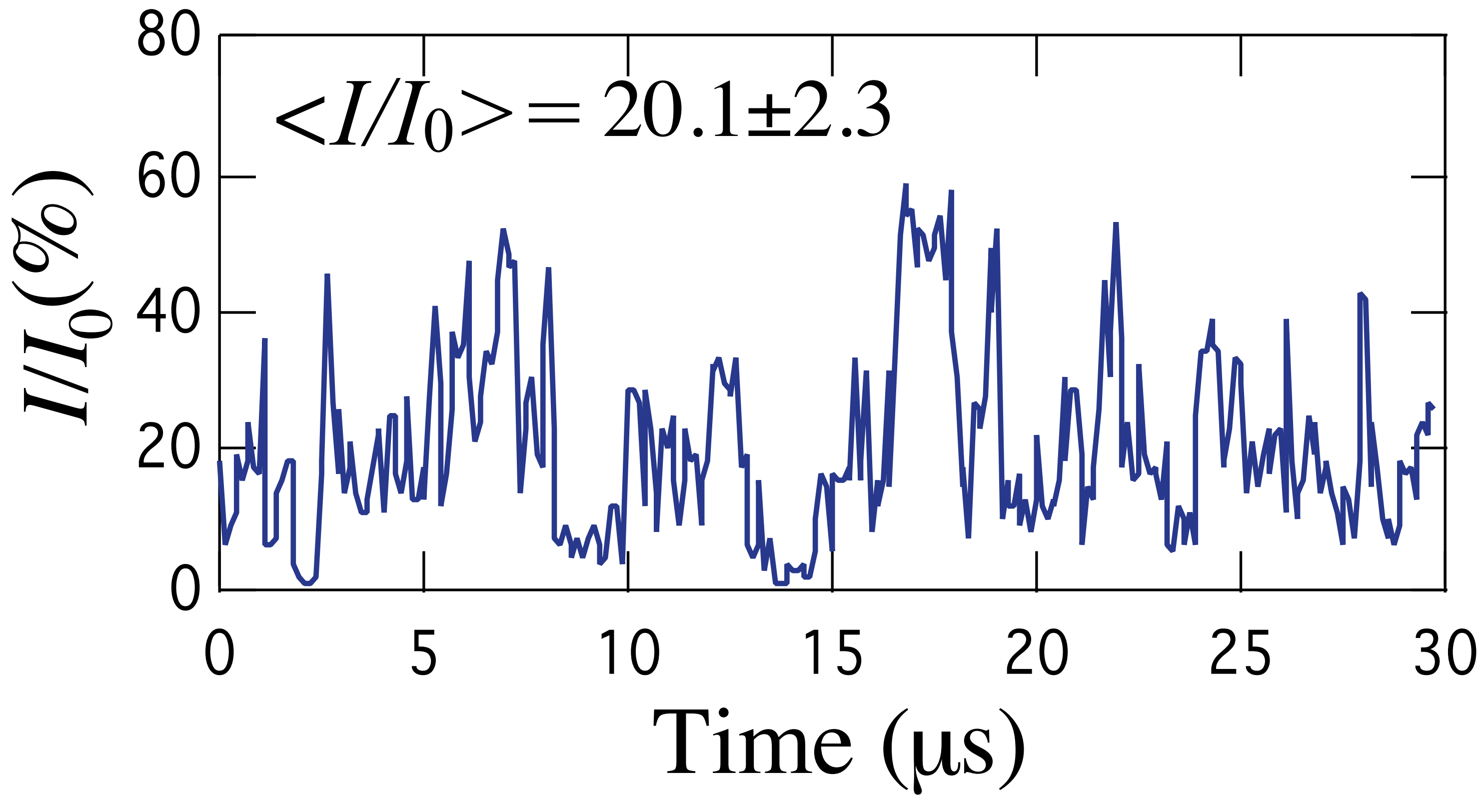
Reduced system (28,000 atoms)

MD simulations of current blockades in MspA

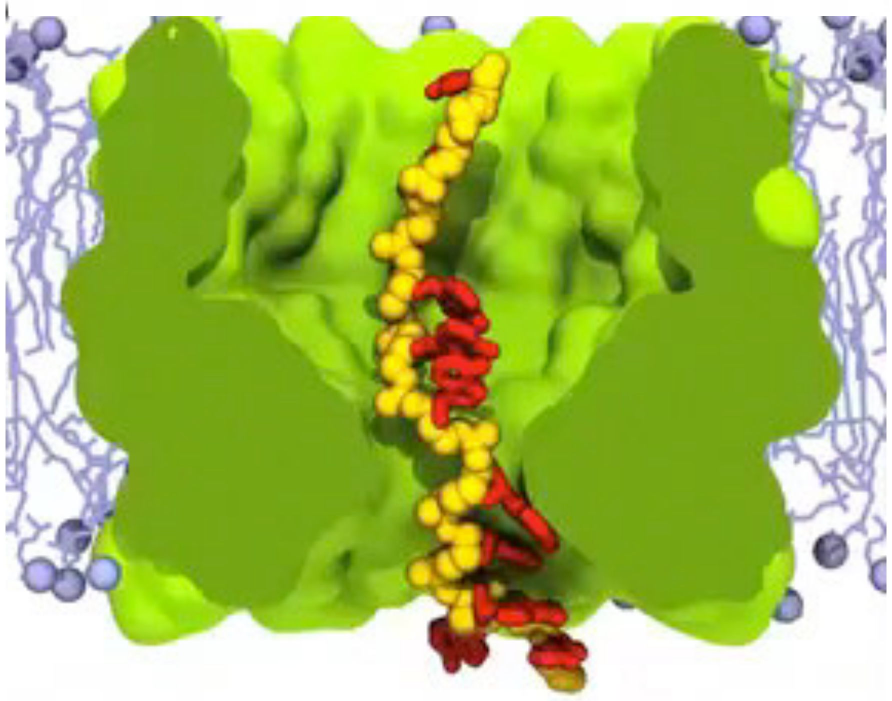
Bhattacharya, Yoo, Aksimentiev, ACS Nano 10, 4644 (2016)

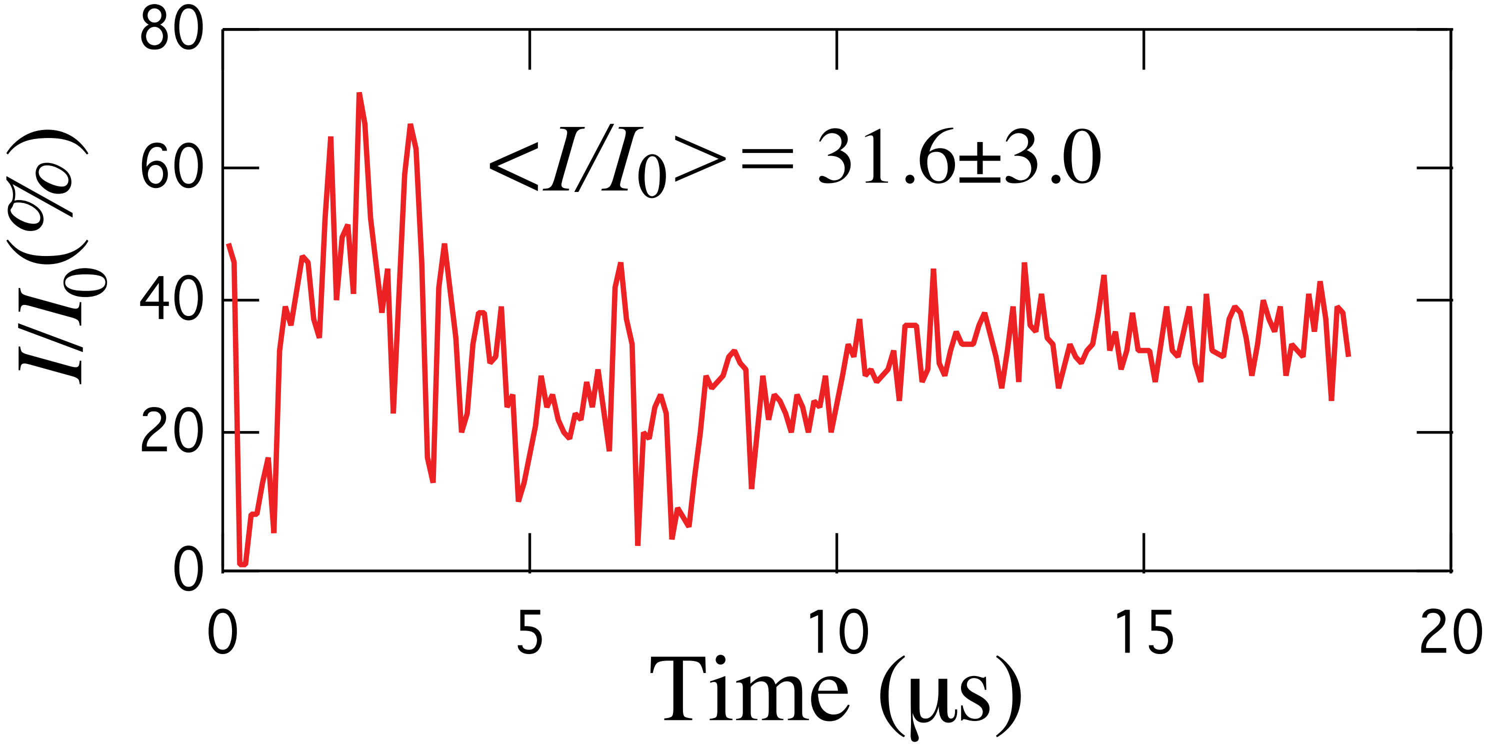
5’-poly(dT)





5’-poly(dA)



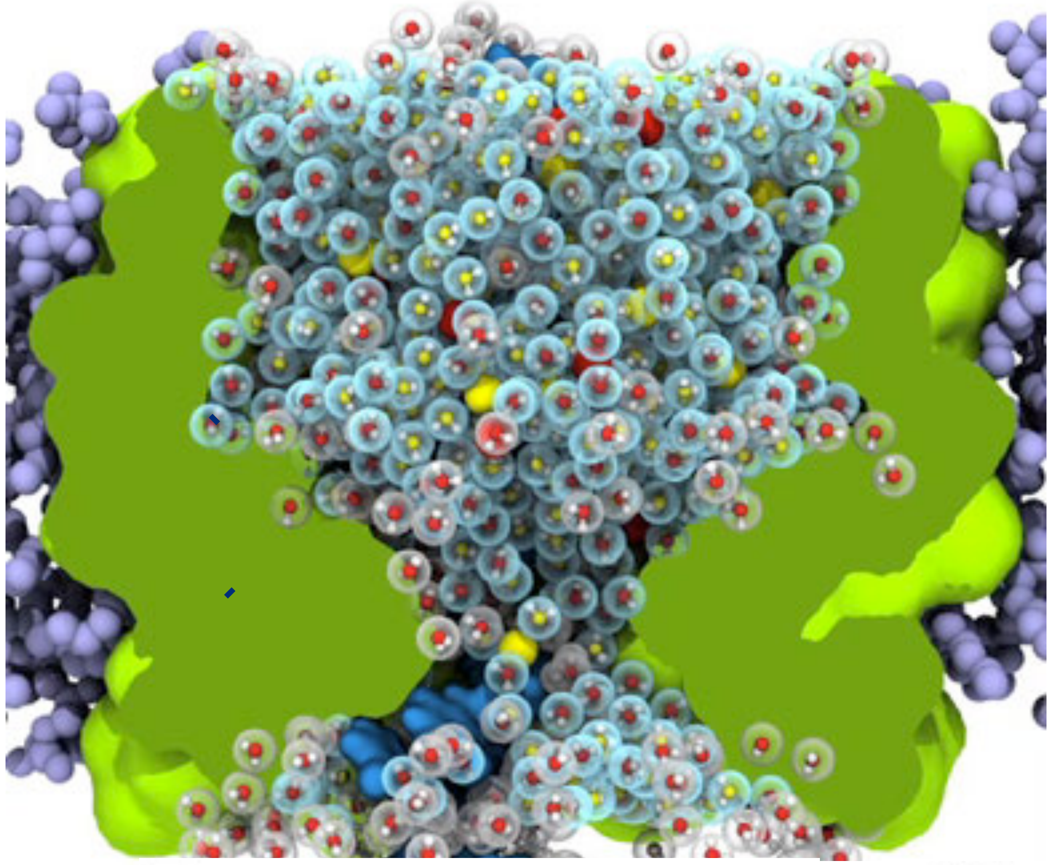


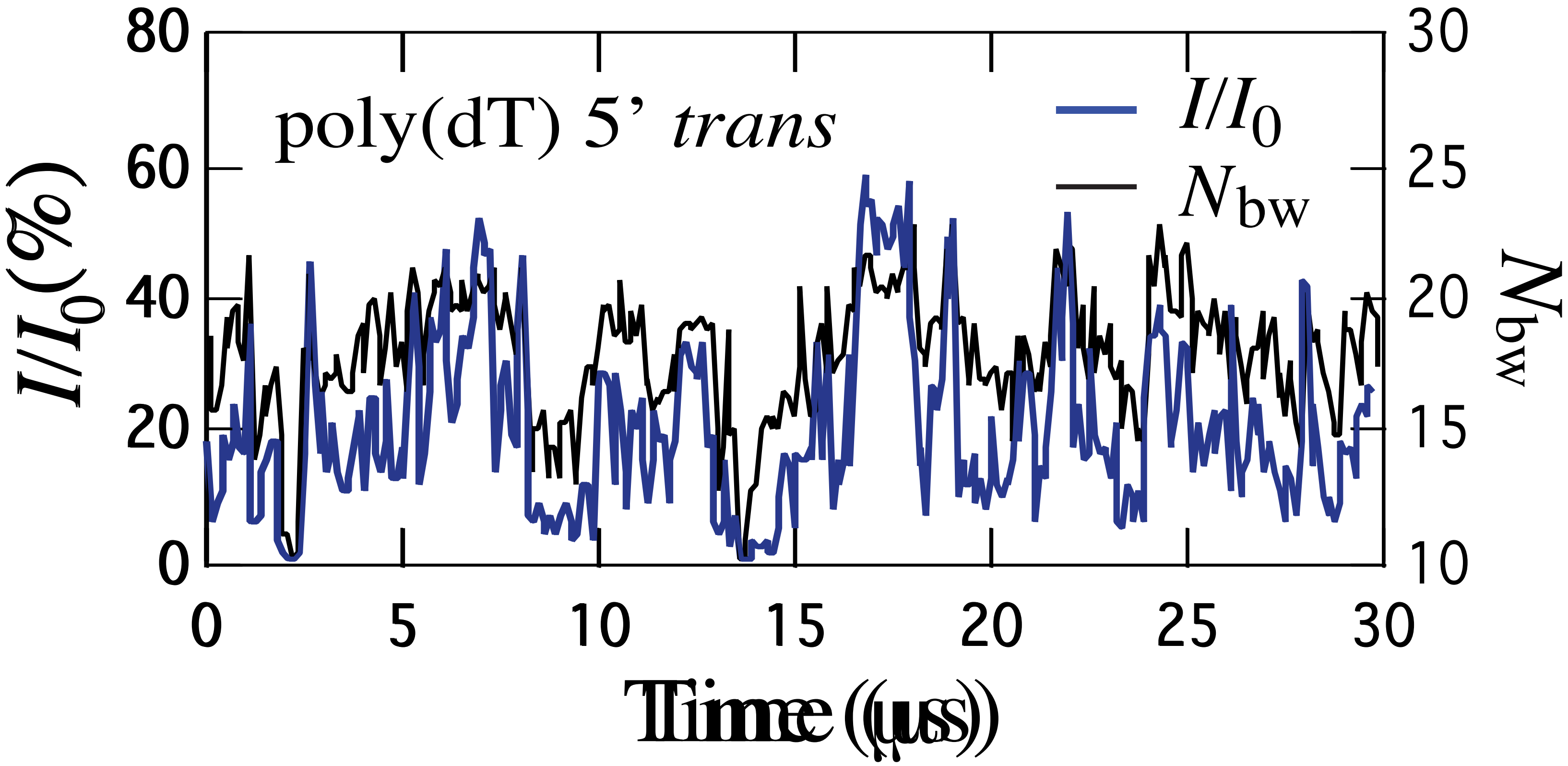
Water mediates DNA sequence recognition

Bhattacharya, Yoo, Aksimentiev, ACS Nano 10, 4644 (2016)

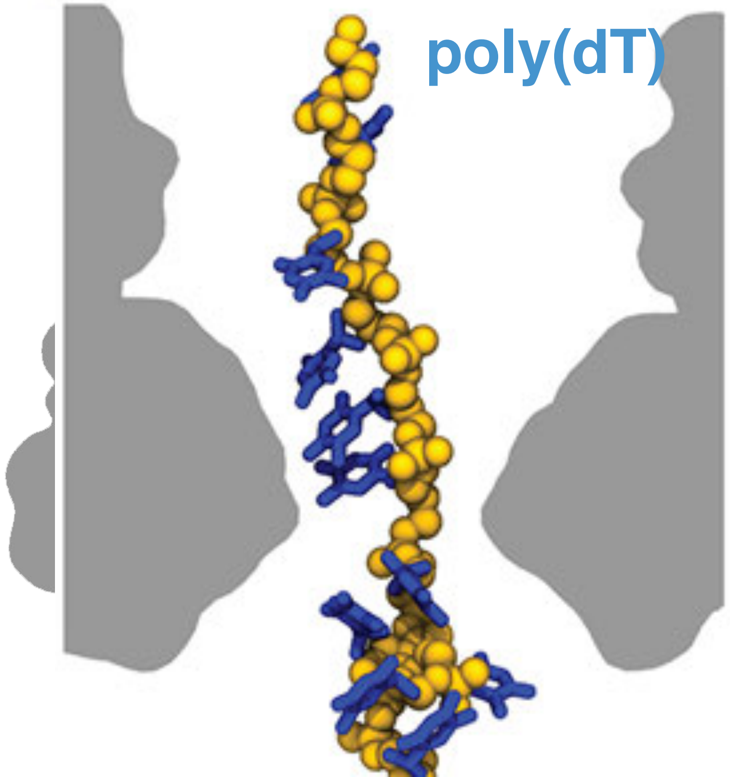
Red: structured water

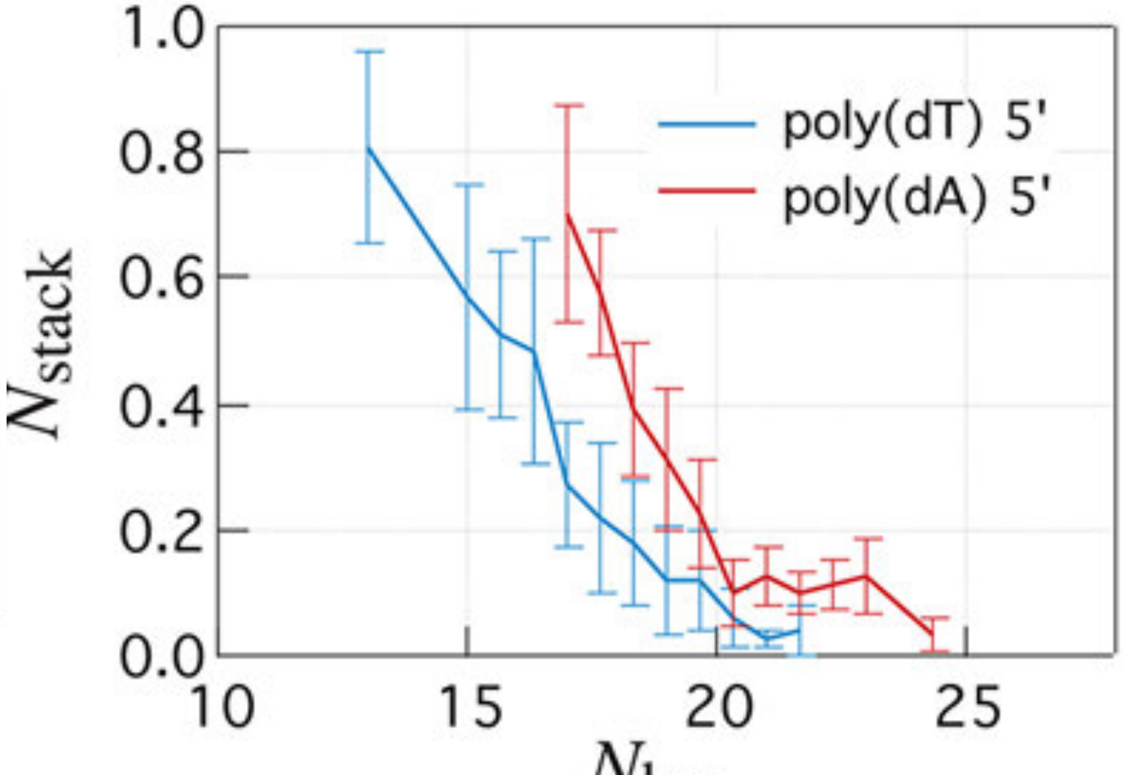
Blue: unstructured water

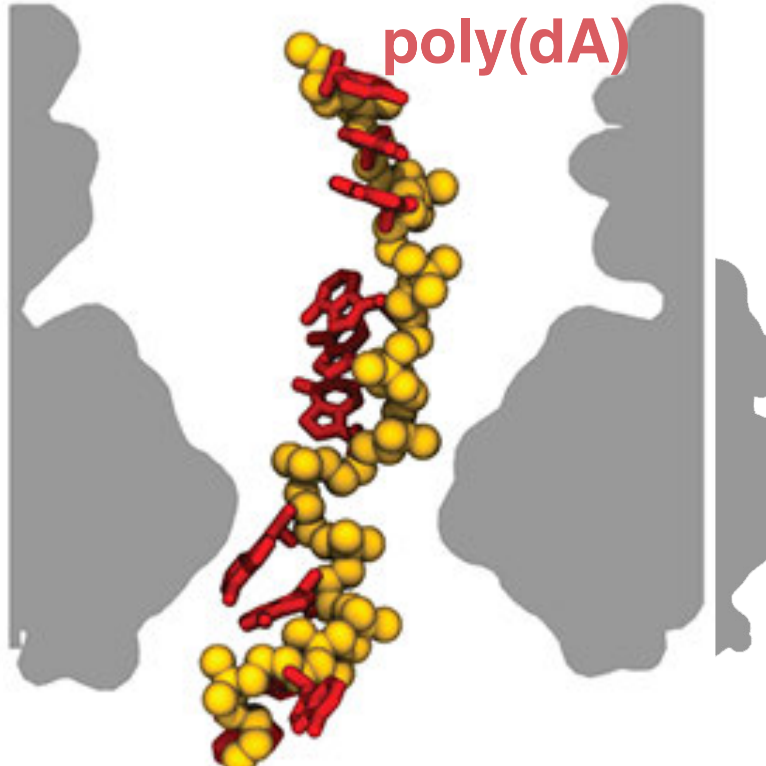




Current increase with the number of water molecules as $N^{1.7}$







MinION (Oxford Nanopore Technologies)

MinION: 800 parallel detection wells

Read length: up to 200,000 nucleotides

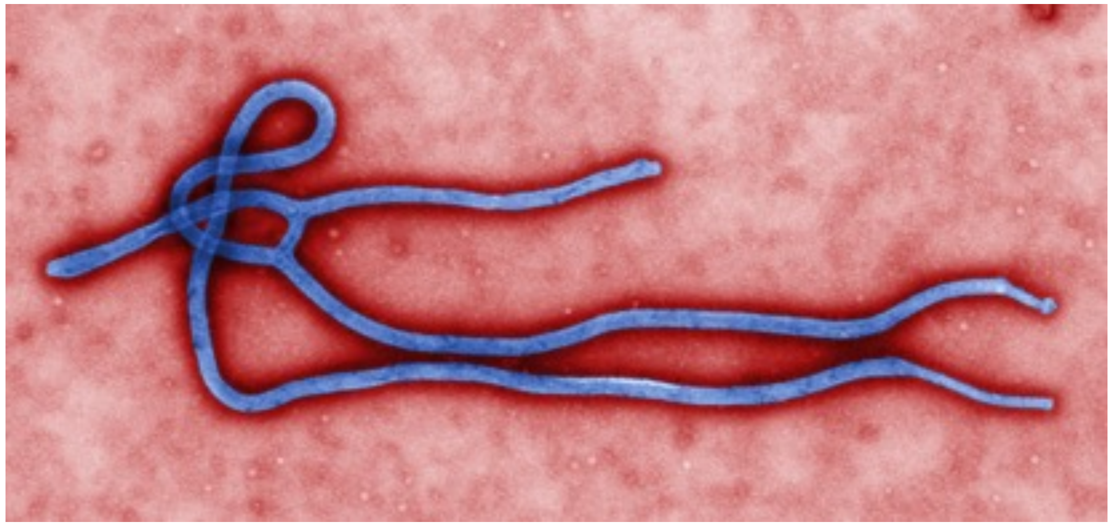
Biological pore (R9, CsgG-derivative)

Helicase motor to control translocation

Direct readout of DNA epigenetic markers

Accuracy: 80 - 98 %

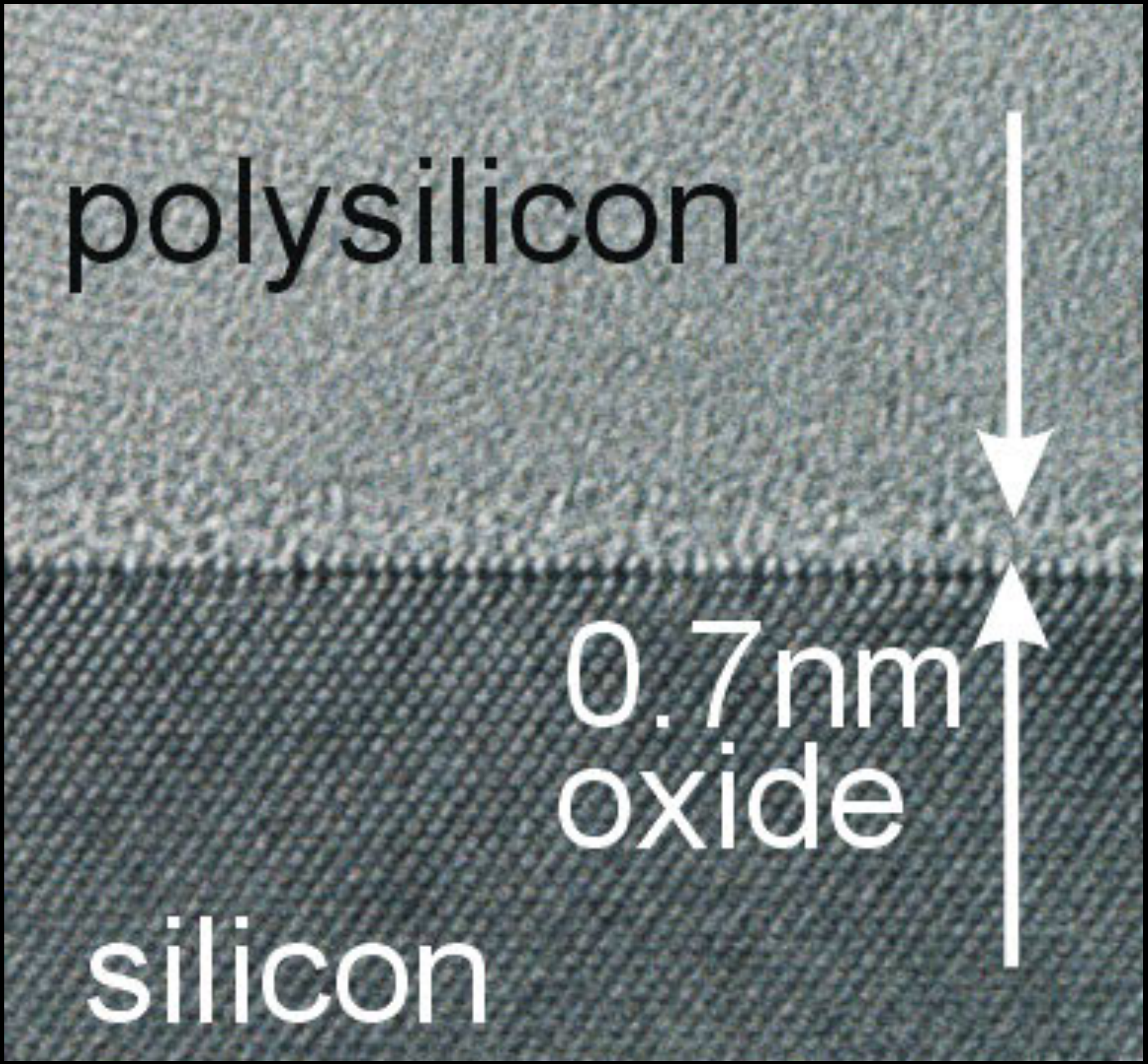
Only 100g, great for *in situ* measurements



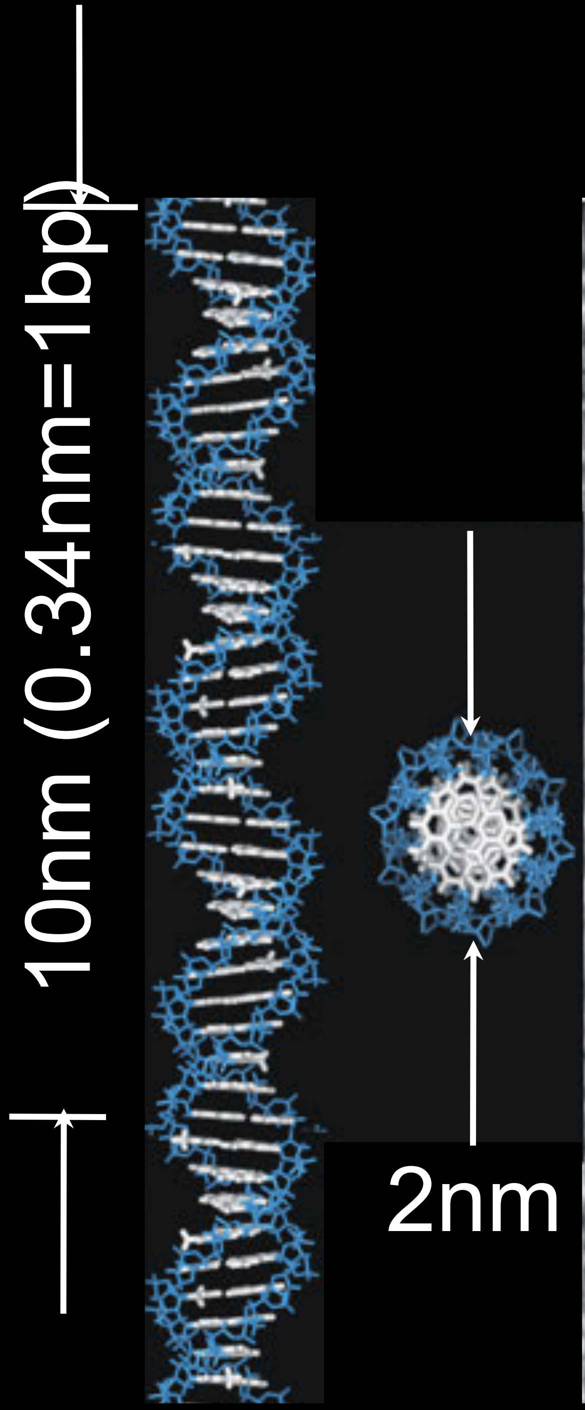
Silicon Nanotechnology for Sequencing DNA

(Back in 2003!)

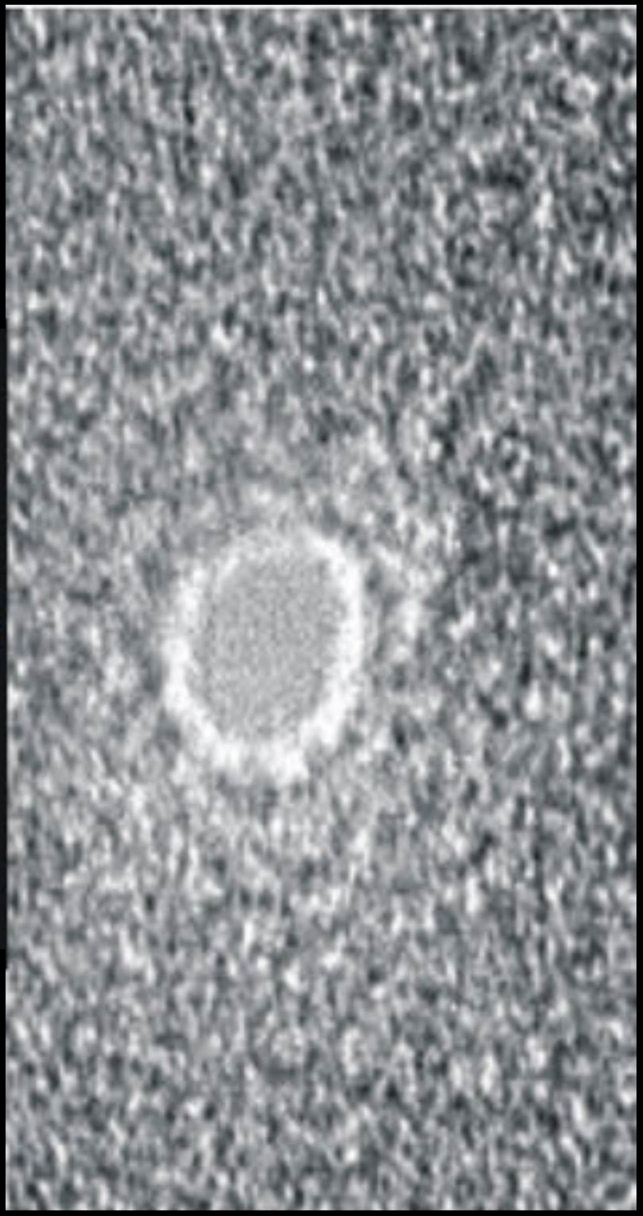
- ultra-thin membranes



TEM X-section through a gate



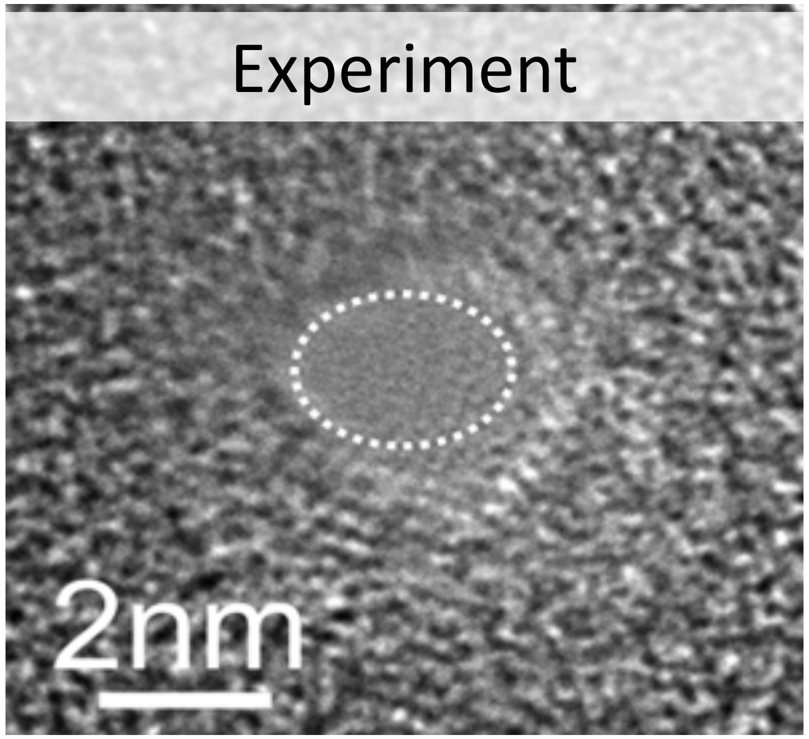
DNA



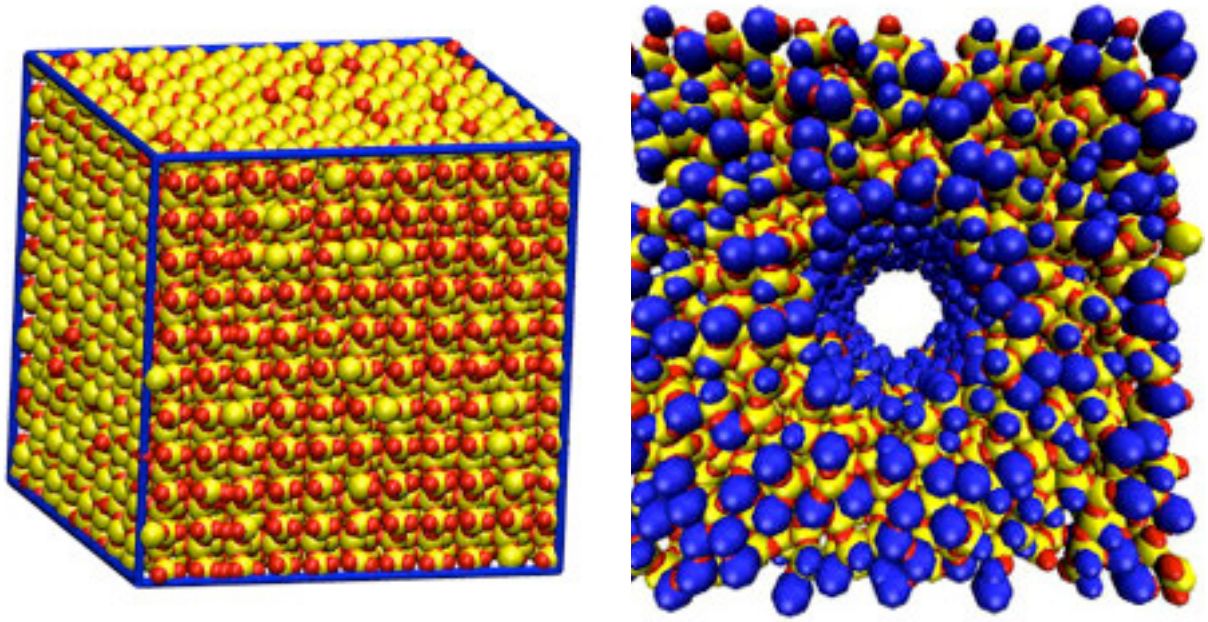
TEM (top-down projection)

T. Sorsch, V. Dimitrov, C. Ho

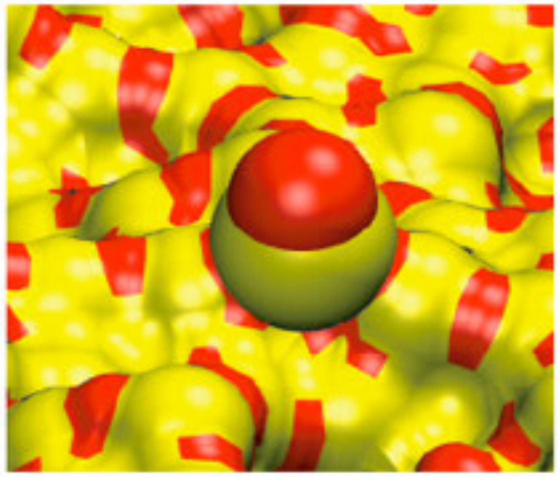
Imaging nanopores using MD



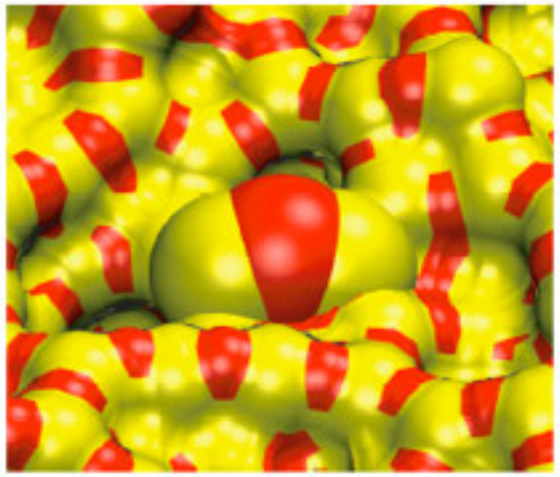
Transmission electron micrograph of a nanopore in Si_3N_4 (Timp Lab, UIUC)



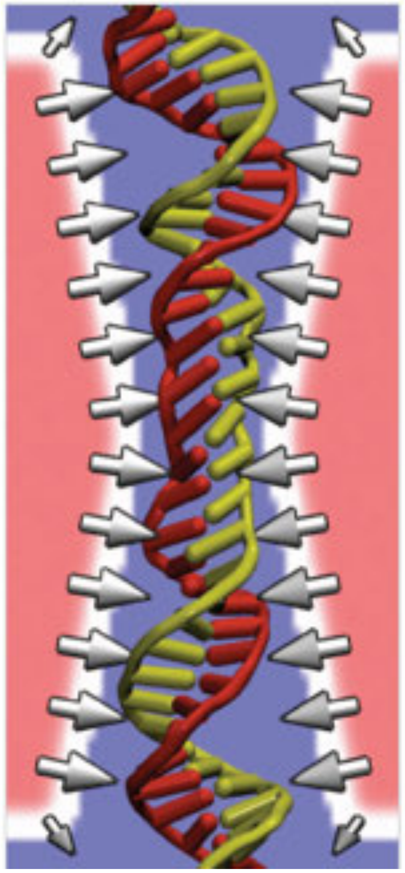
J. Phys. Chem. B 2006, 110, 21497



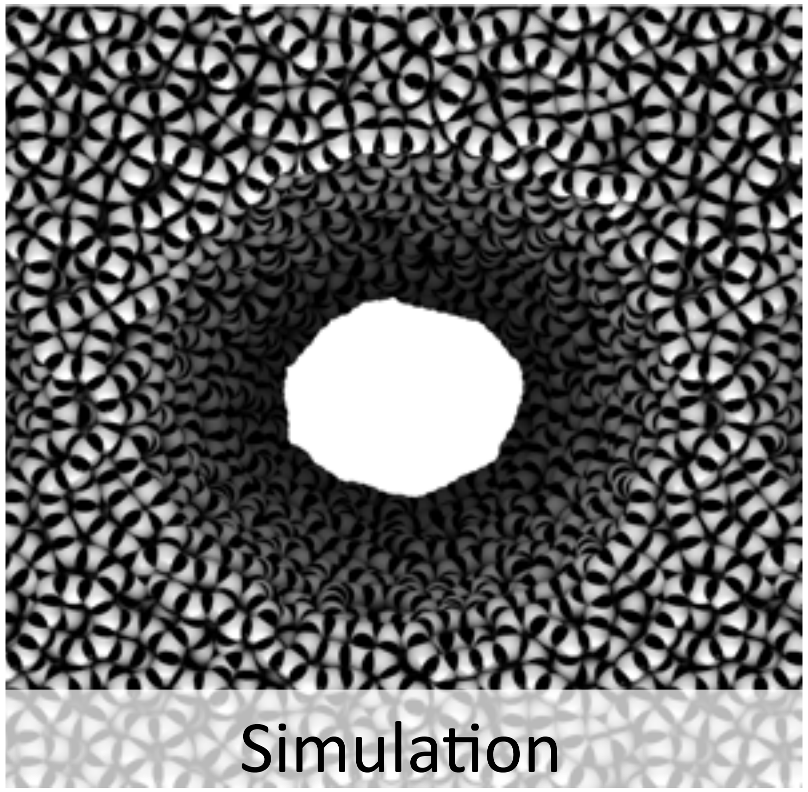
Dangling Oxygen



Non-Dangling Oxygen

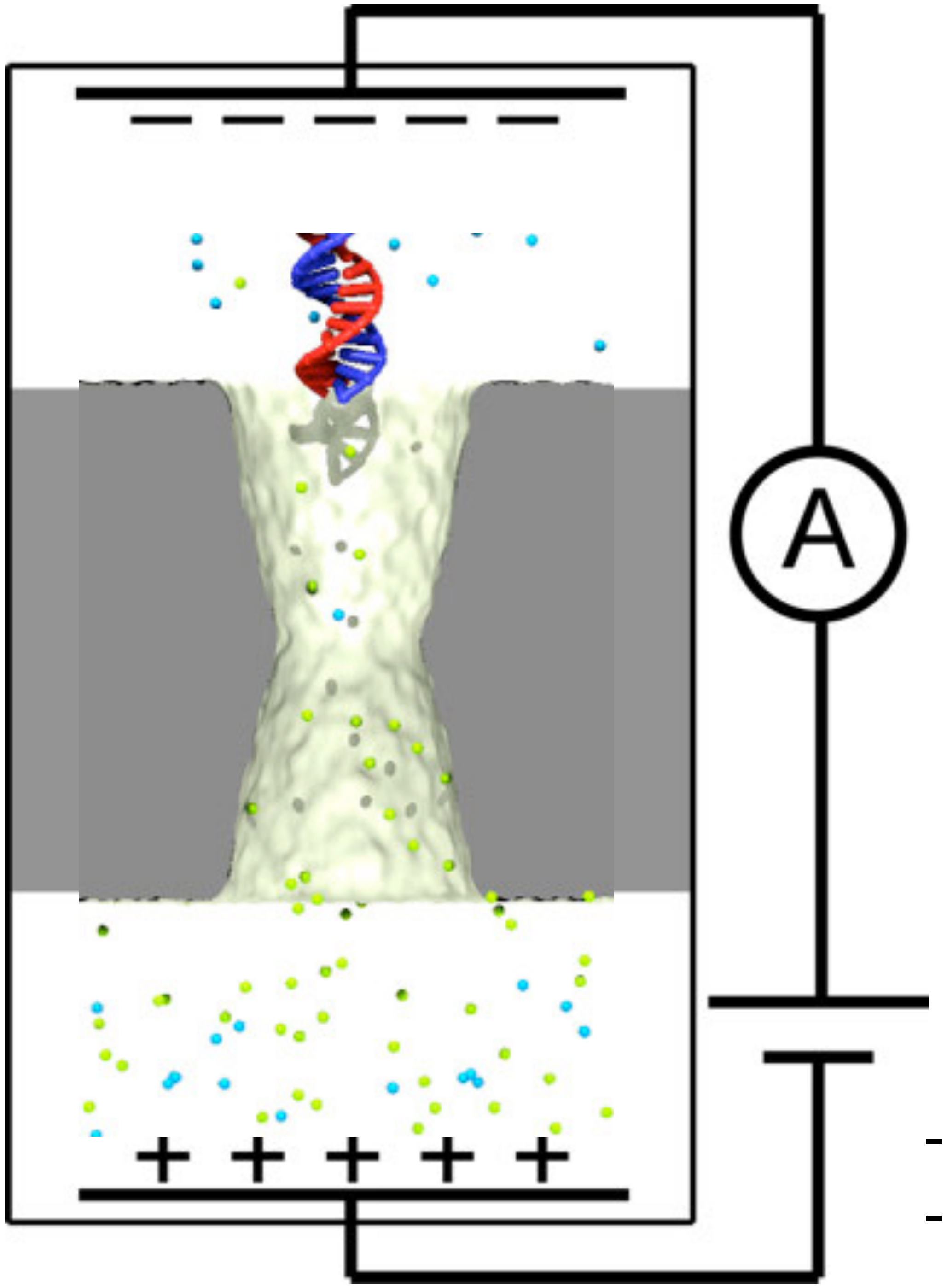


Grid based surface potential

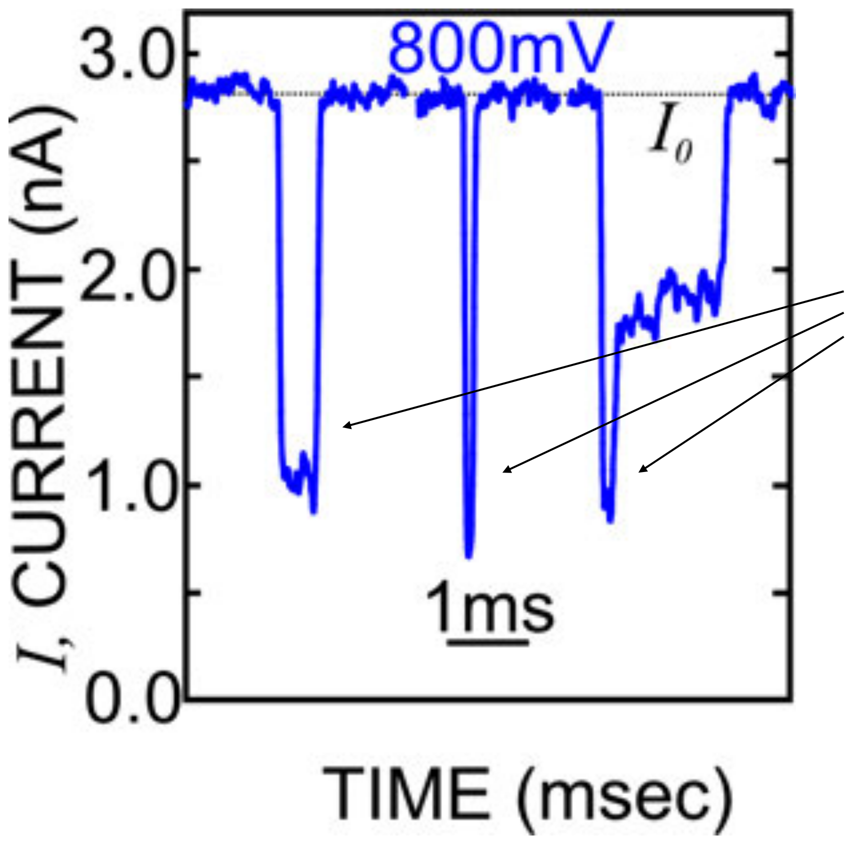


All atom model of a nanopore in Si_3N_4 built by the Aksimentiev group

DNA electrophoresis through a nanopore



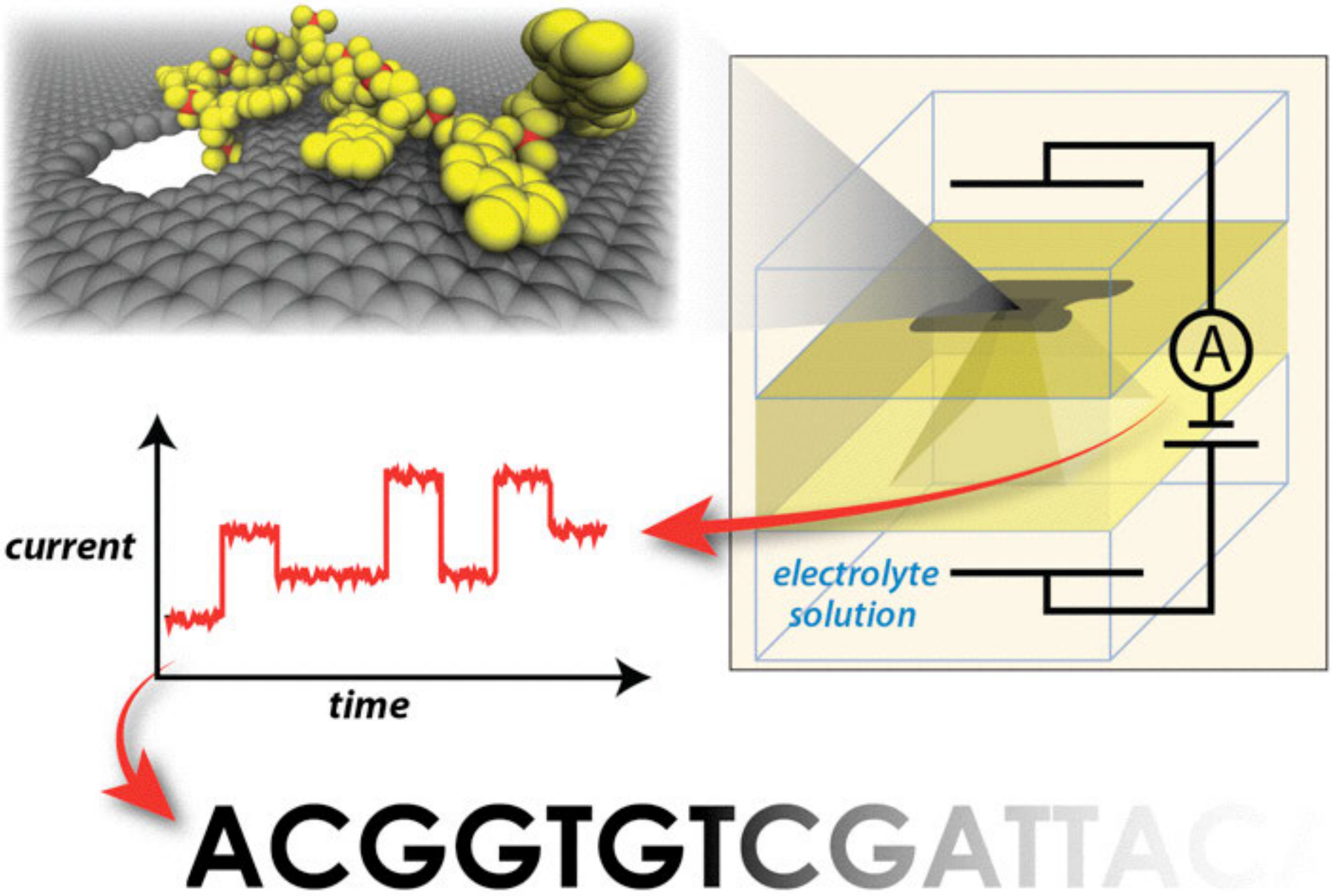
Ionic current through pore measured



current transients associated with passage of dsDNA

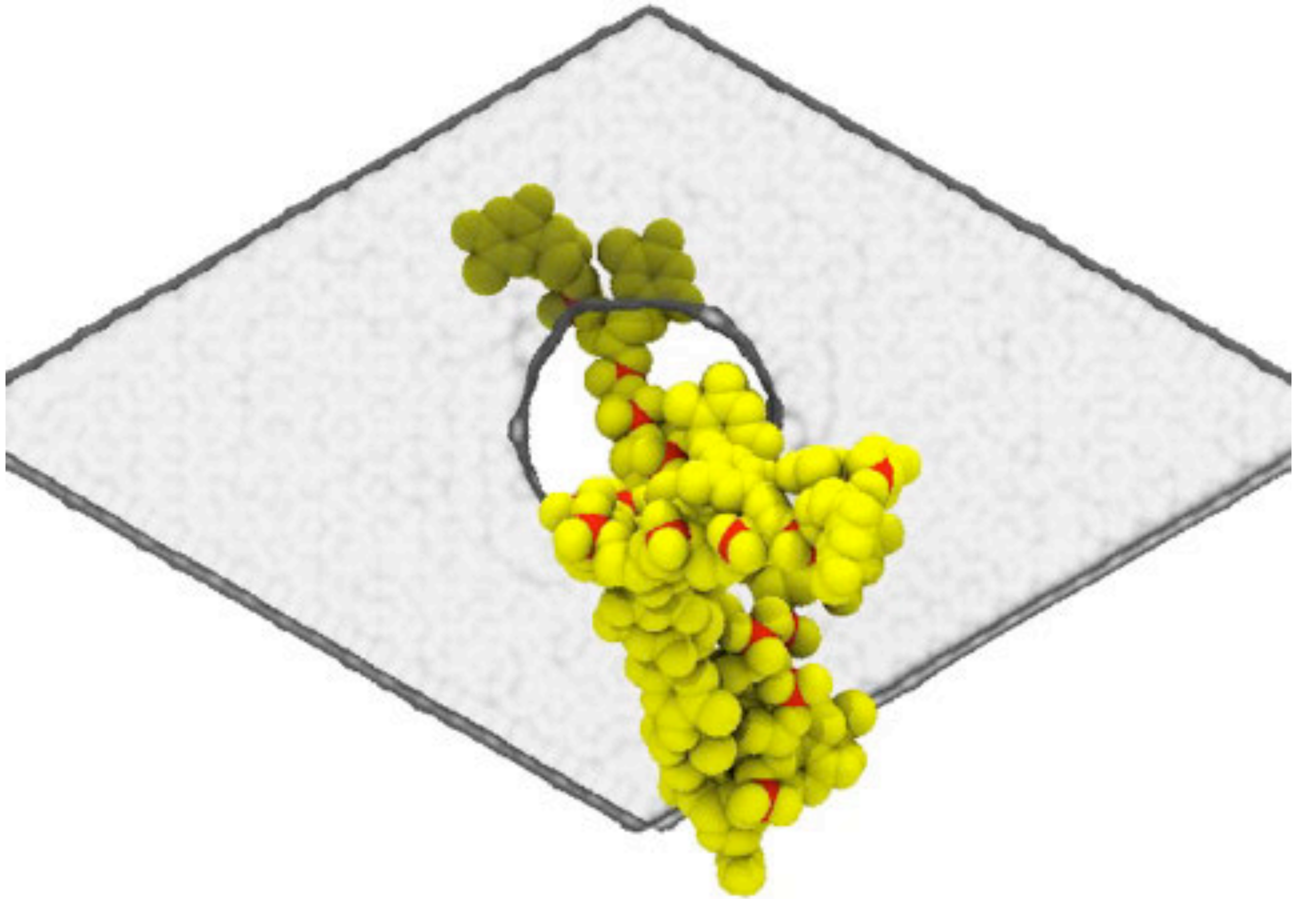
- Isolates 1nm^3 of volume
- Automatic loading and reloading
- Highly processive, single-file transport
- Compatible with several detection schemes
- No limit on the read length

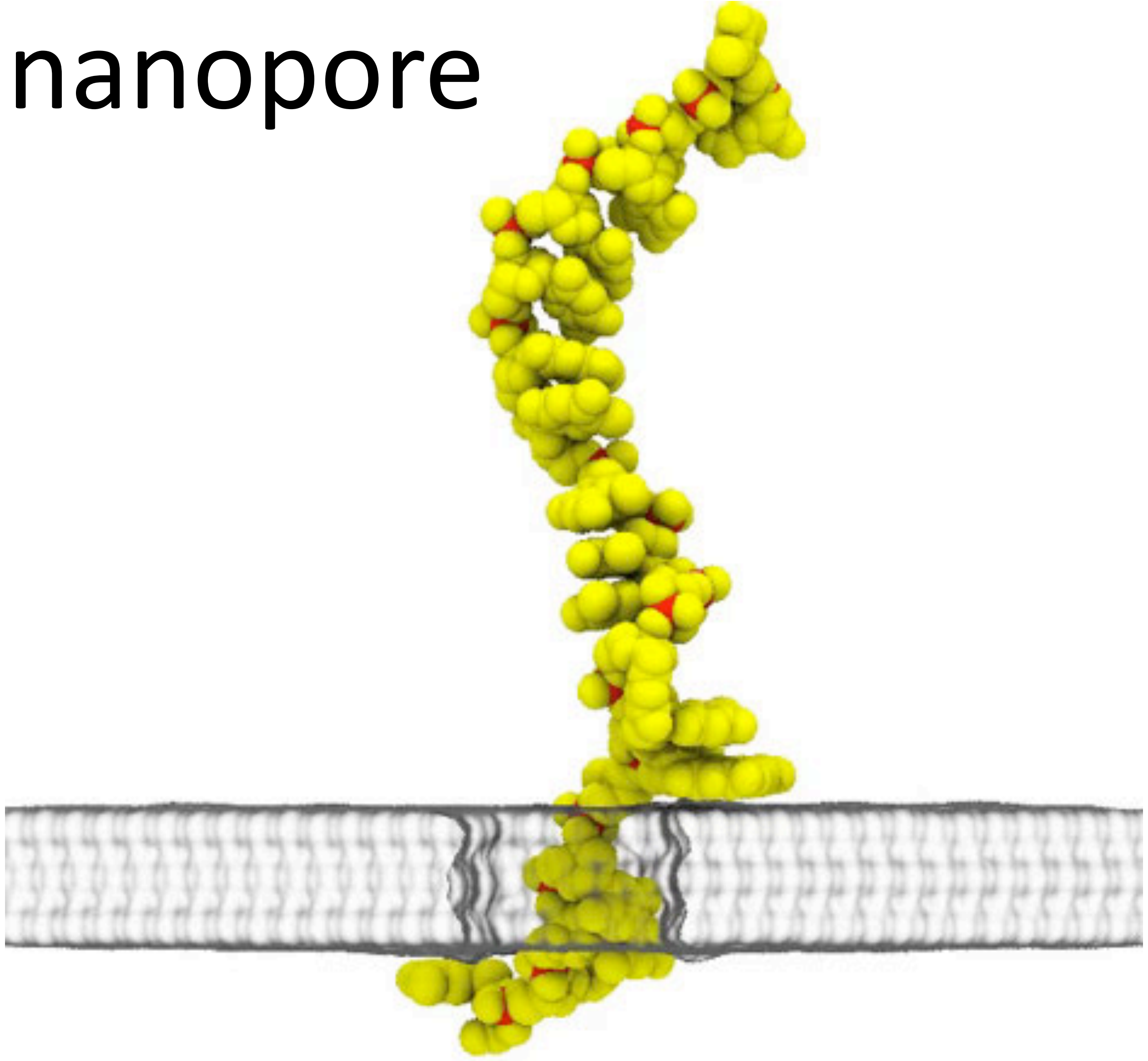
Graphene Nanopores



Simulation of DNA translocation through graphene nanopore

Top view





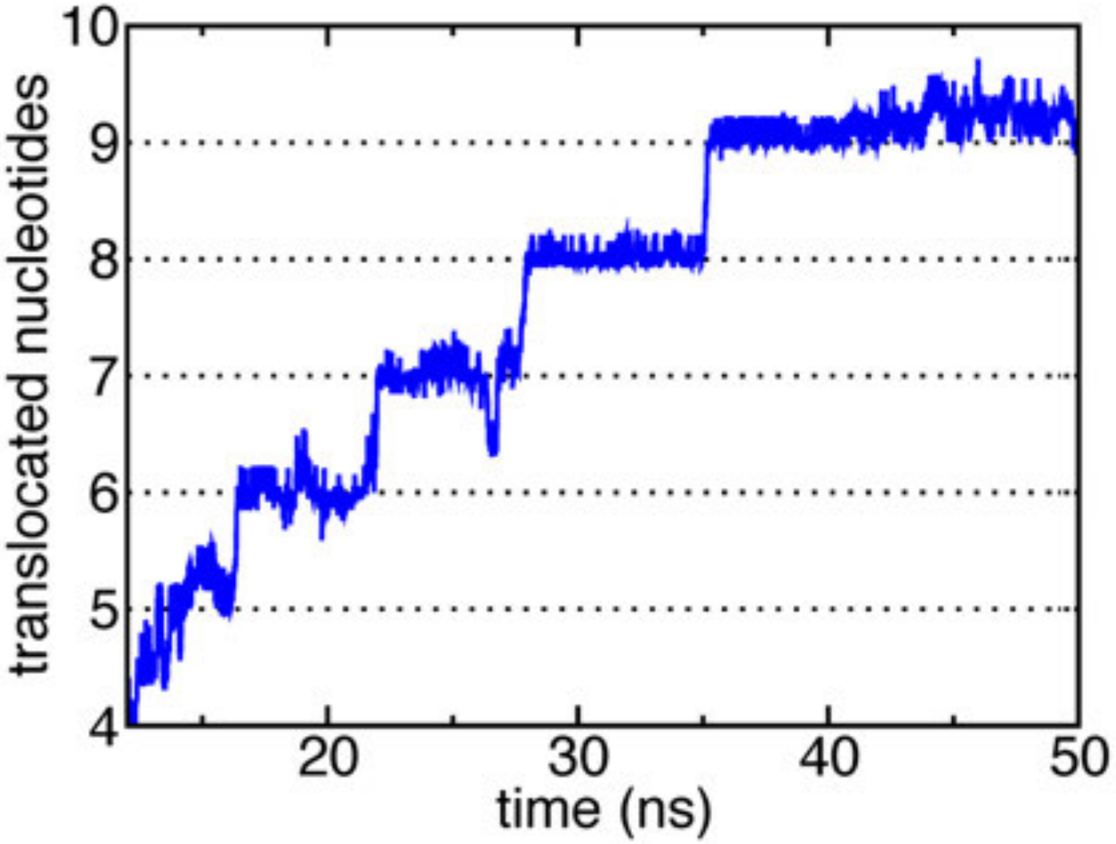
Side view

14-A diameter pore (surface-to surface);
3-layer graphite;
poly(dT)$_{20}$; 500 mV bias

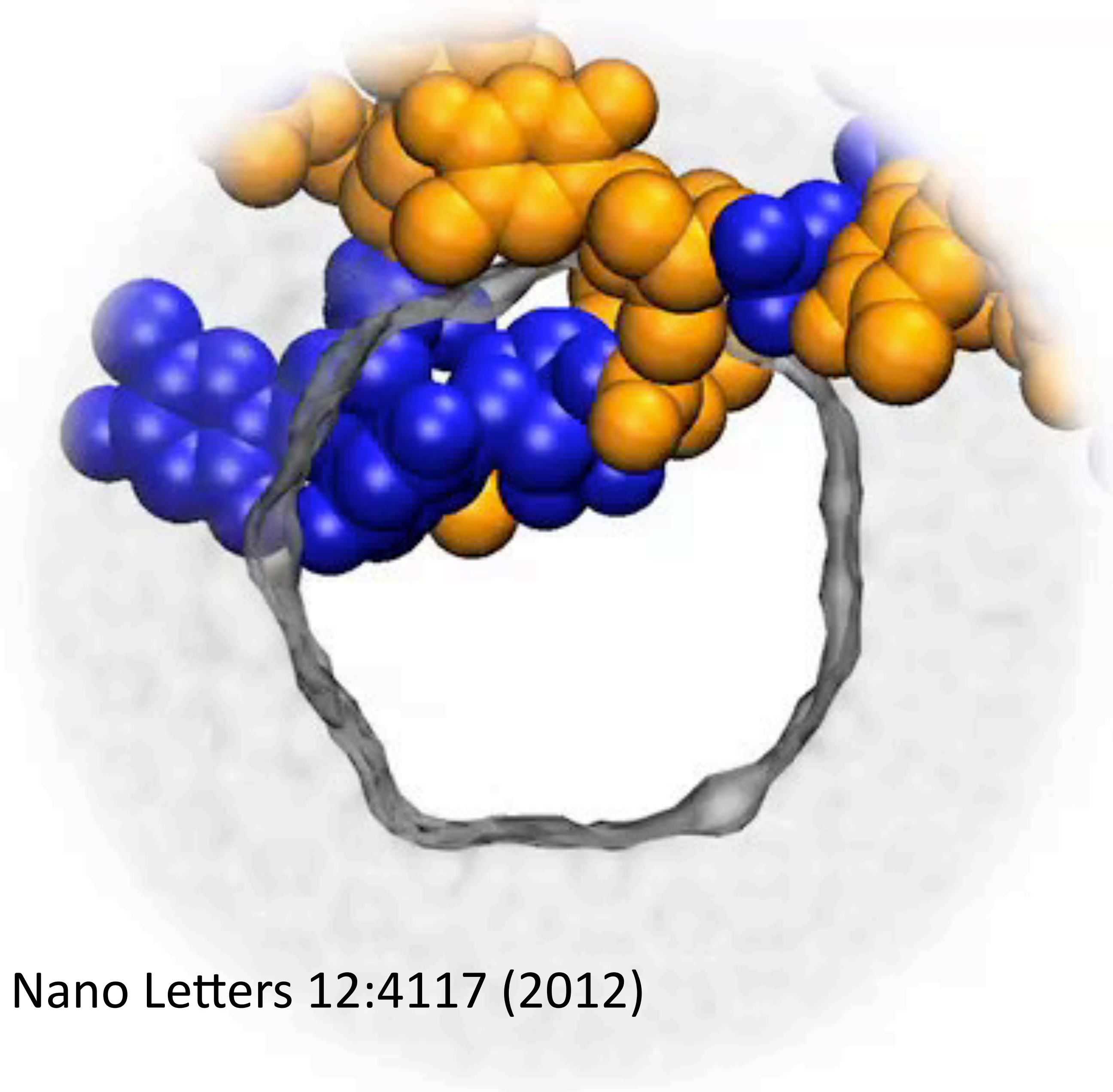
Nano Letters 12:4117 (2012)

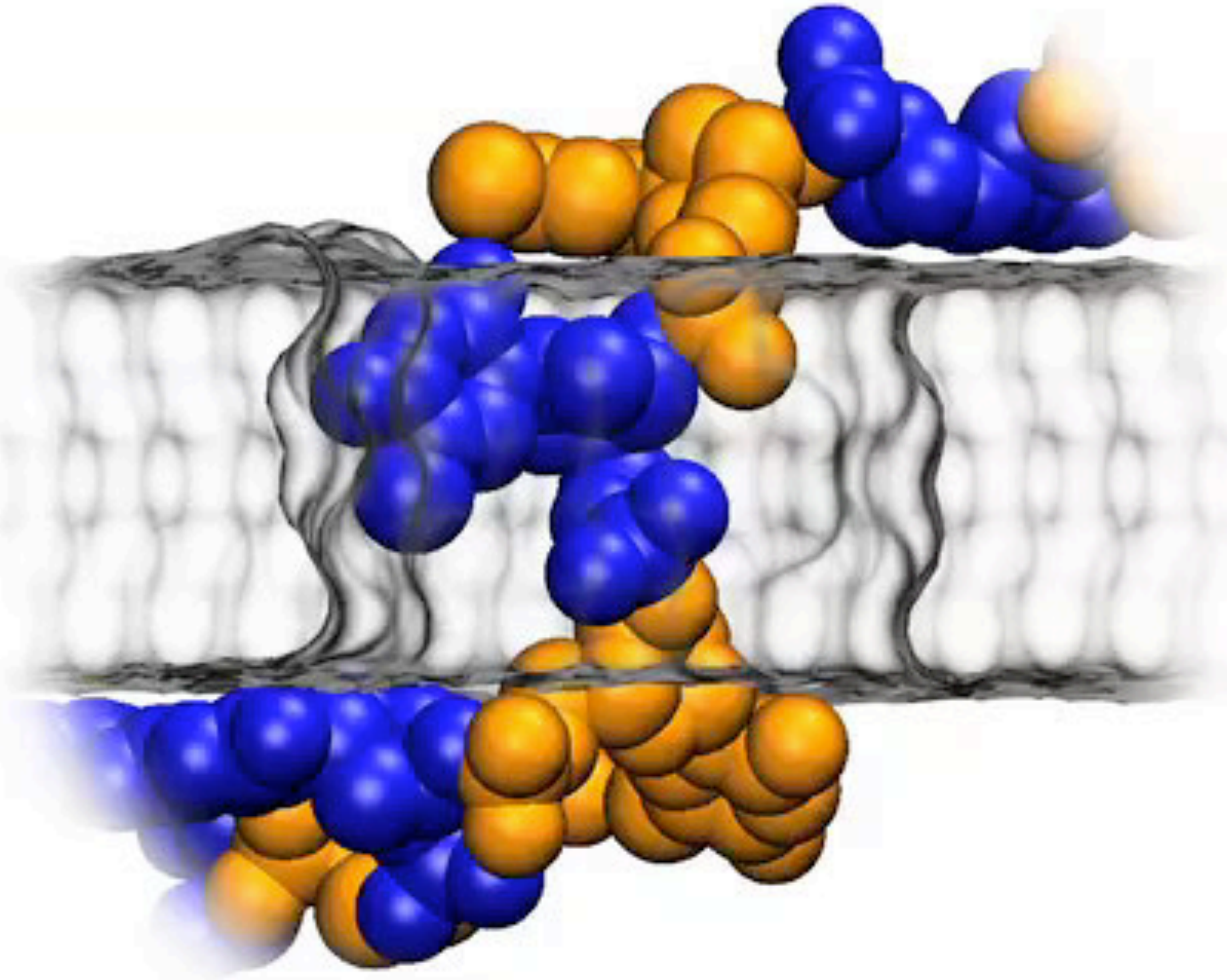
DNA transport is stepwise

14-A diameter pore (surface-to surface);
3-layer graphite;
poly(dT)$_{20}$; 500 mV bias



DNA-graphene interactions
act as a stepping motor!

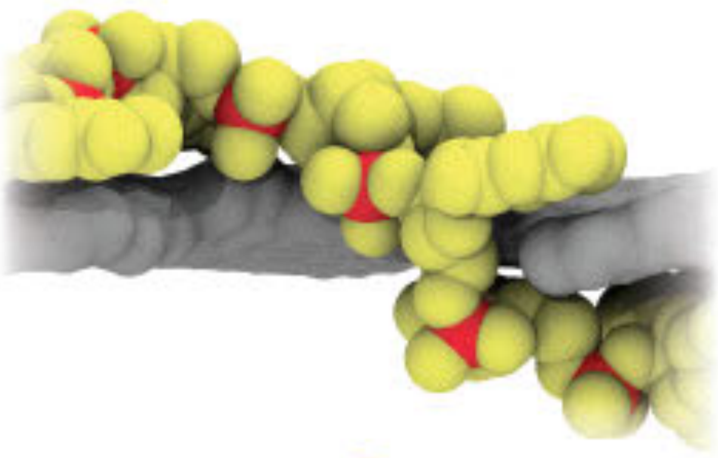




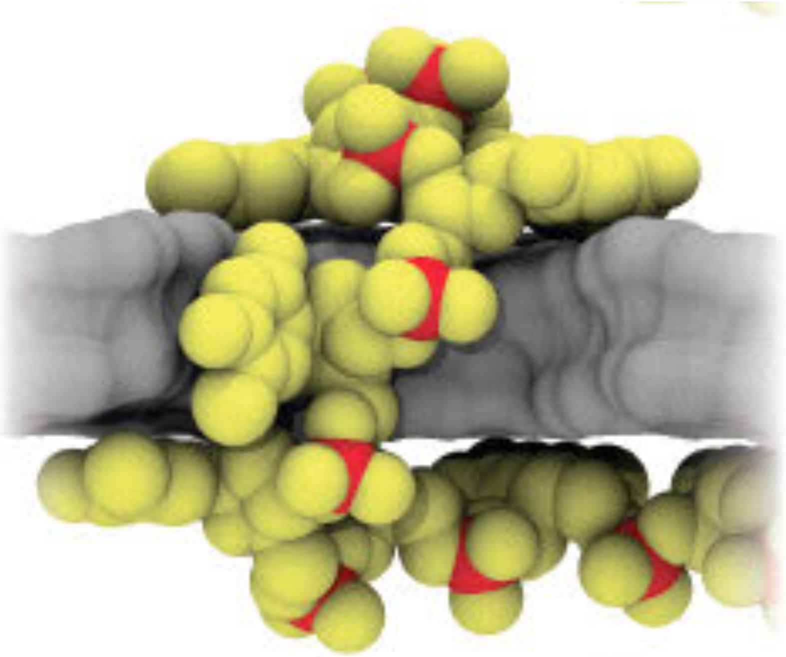
Nano Letters 12:4117 (2012)

Can ionic current blockades can reveal the DNA sequence?

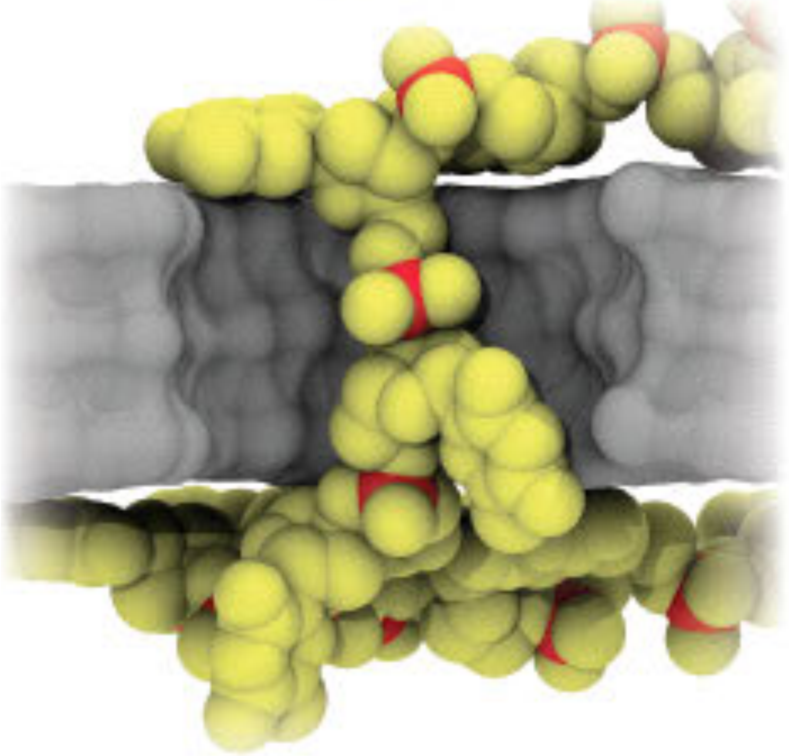
single layer

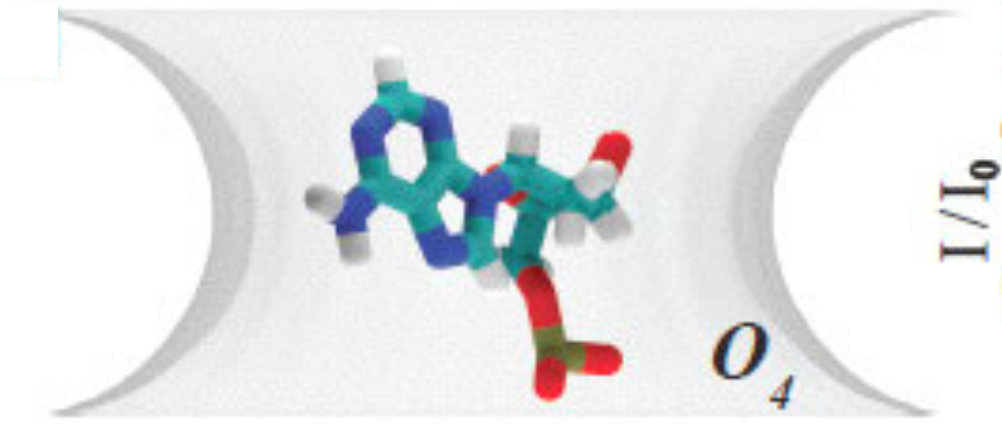


double layer

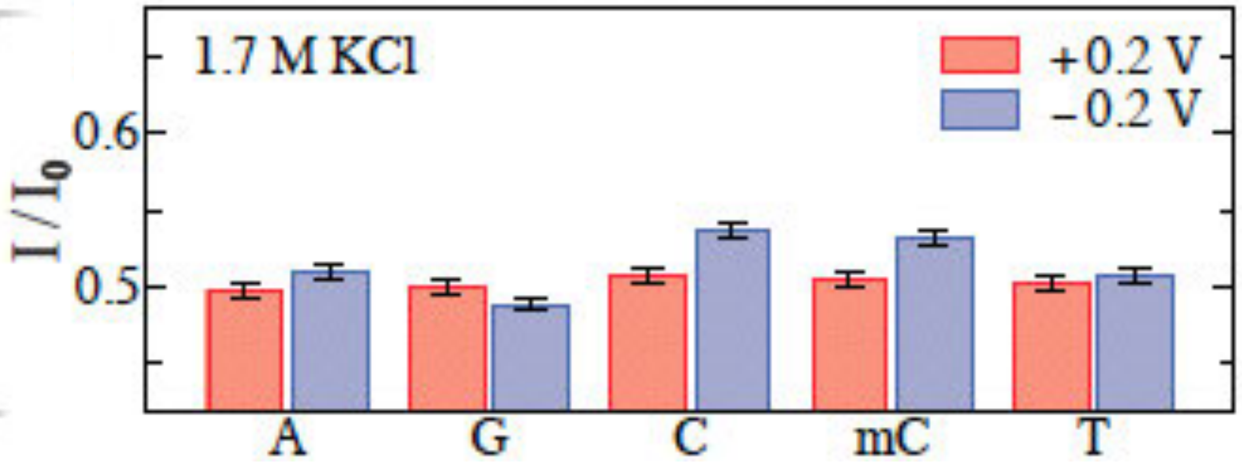


tripple layer





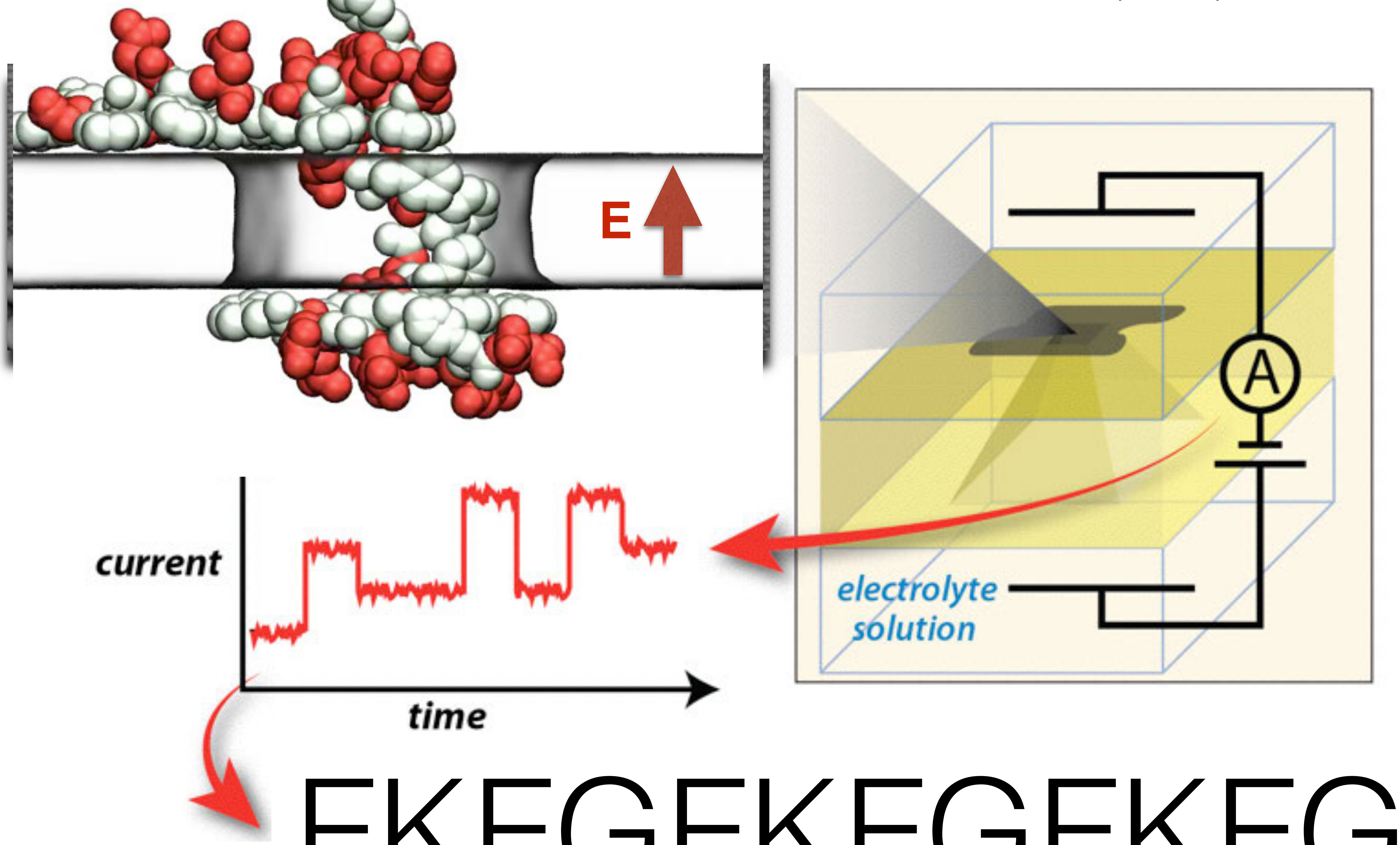
Nano Letters 12:4117 (2012)



Atomic-Resolution Brownian Dynamics simulations of ionic current blockades in graphene nanopores

Sequencing proteins using graphene nanopores?

Wilson et al., Adv. Func. Mat. 26:4830-4838 (2016)



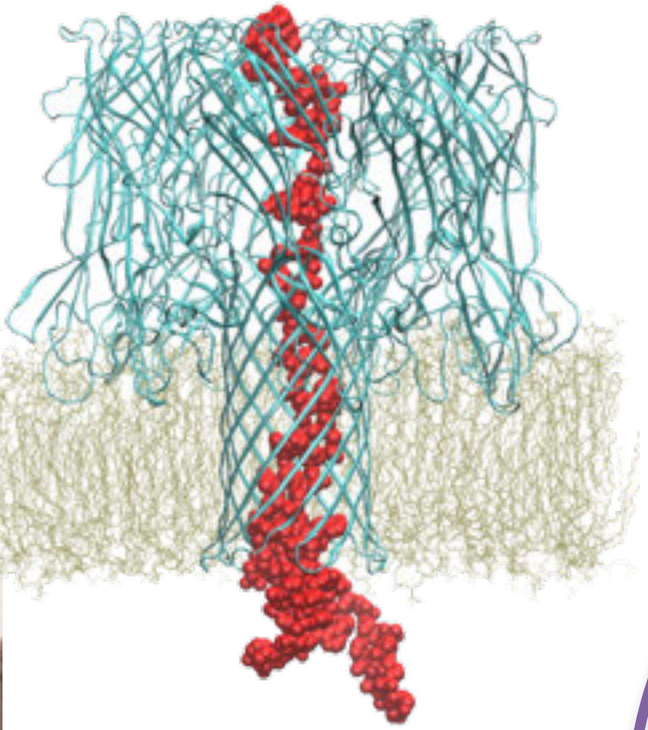
FKFGFKFGFKFG

Modeling Nanopores for Sequencing DNA

Biological Nanopores

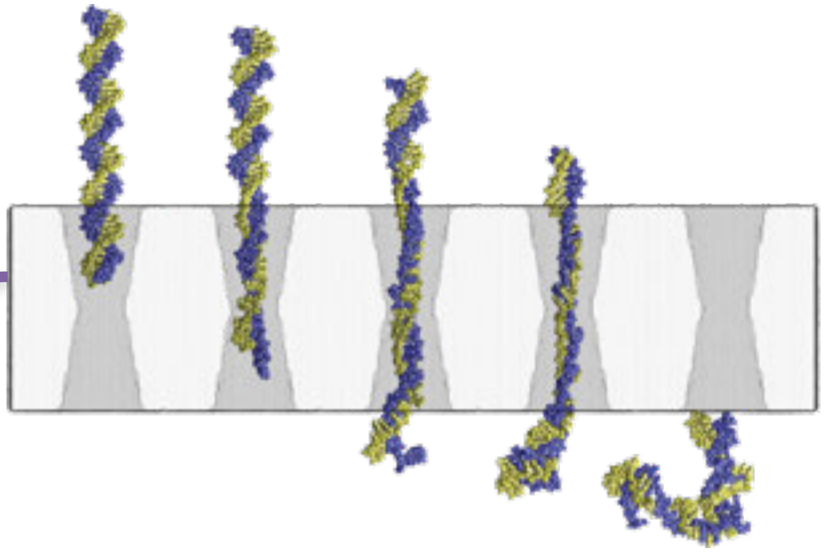
α-Hemolysin

Tutorial Difficulty: Easy



Silicone Nitride

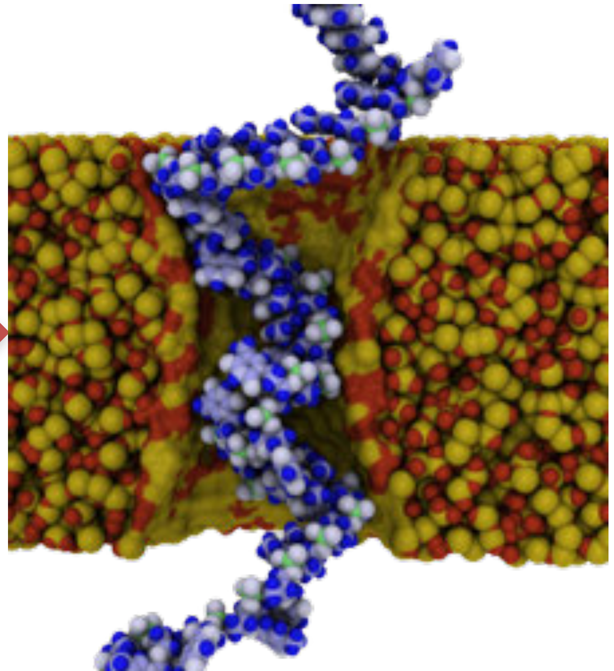
Tutorial Difficulty: Medium



Manish Shankla

Silica

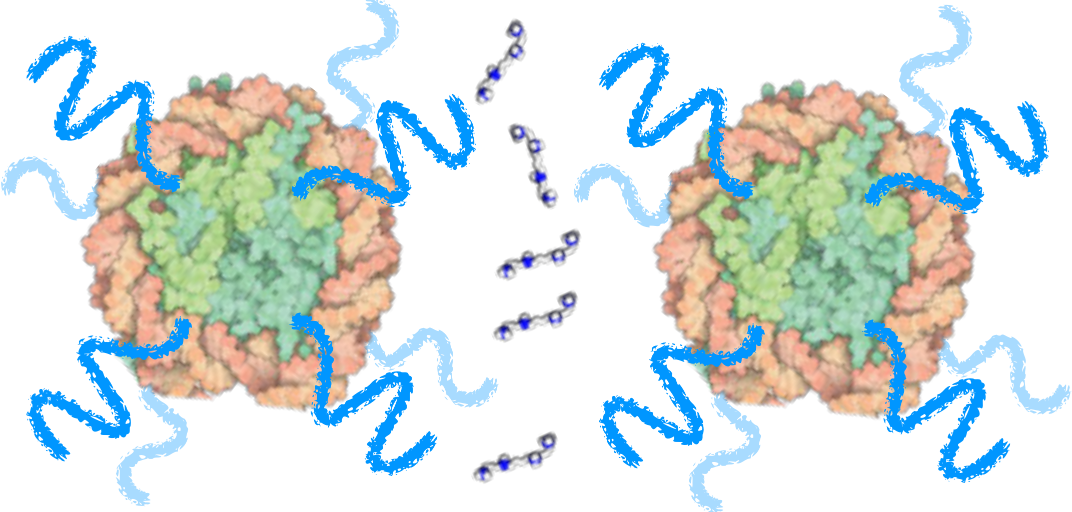
Tutorial Difficulty: NIGHTMARE



Solid-State Nanopores

What is the next step?

- Biology is understood



- Means to detect a problem developed

- How are we going to fix the problem?

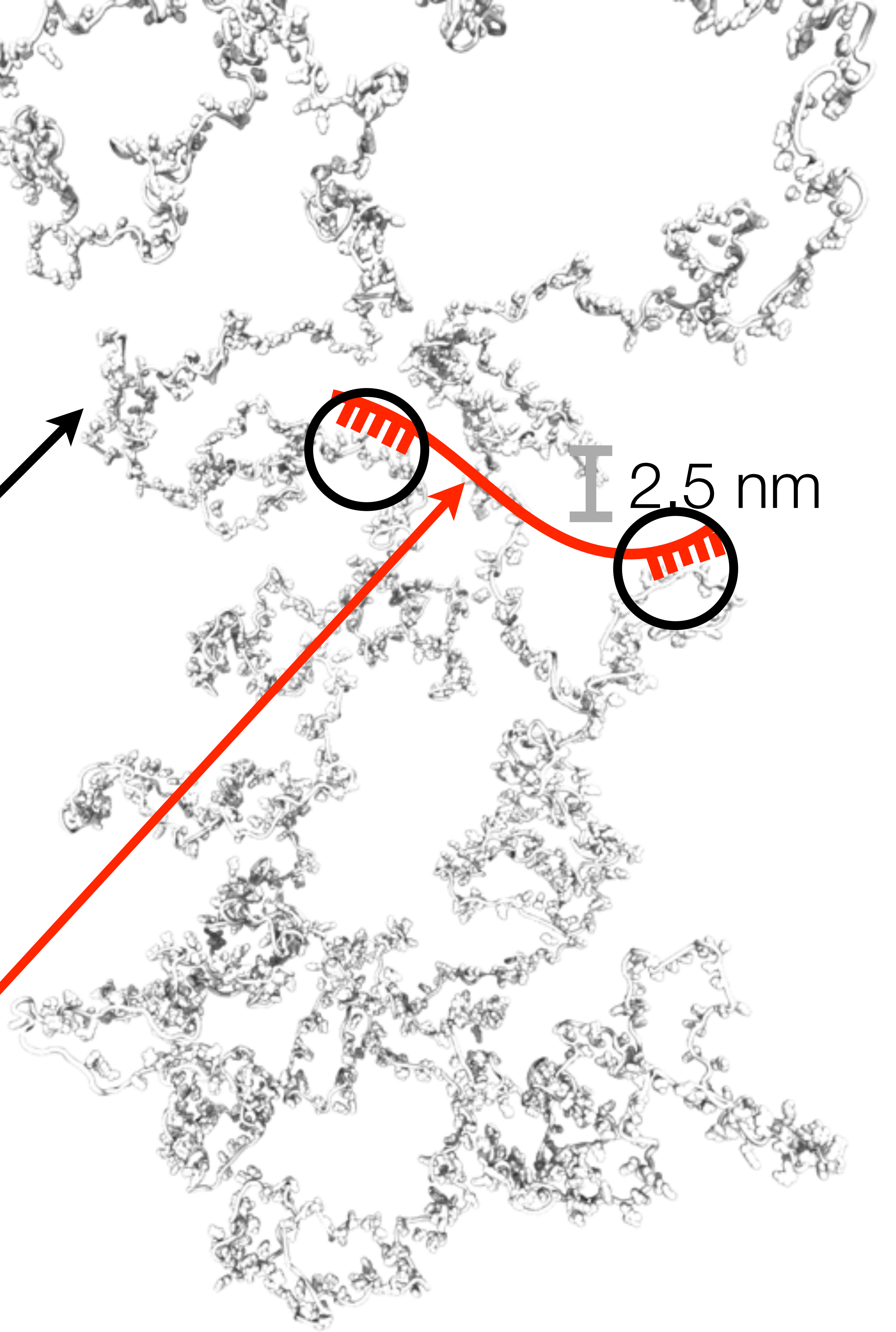
Target ONLY diseased cells

Prevent degradation before reaching target

Make sure the drug enters a cell

DNA origami

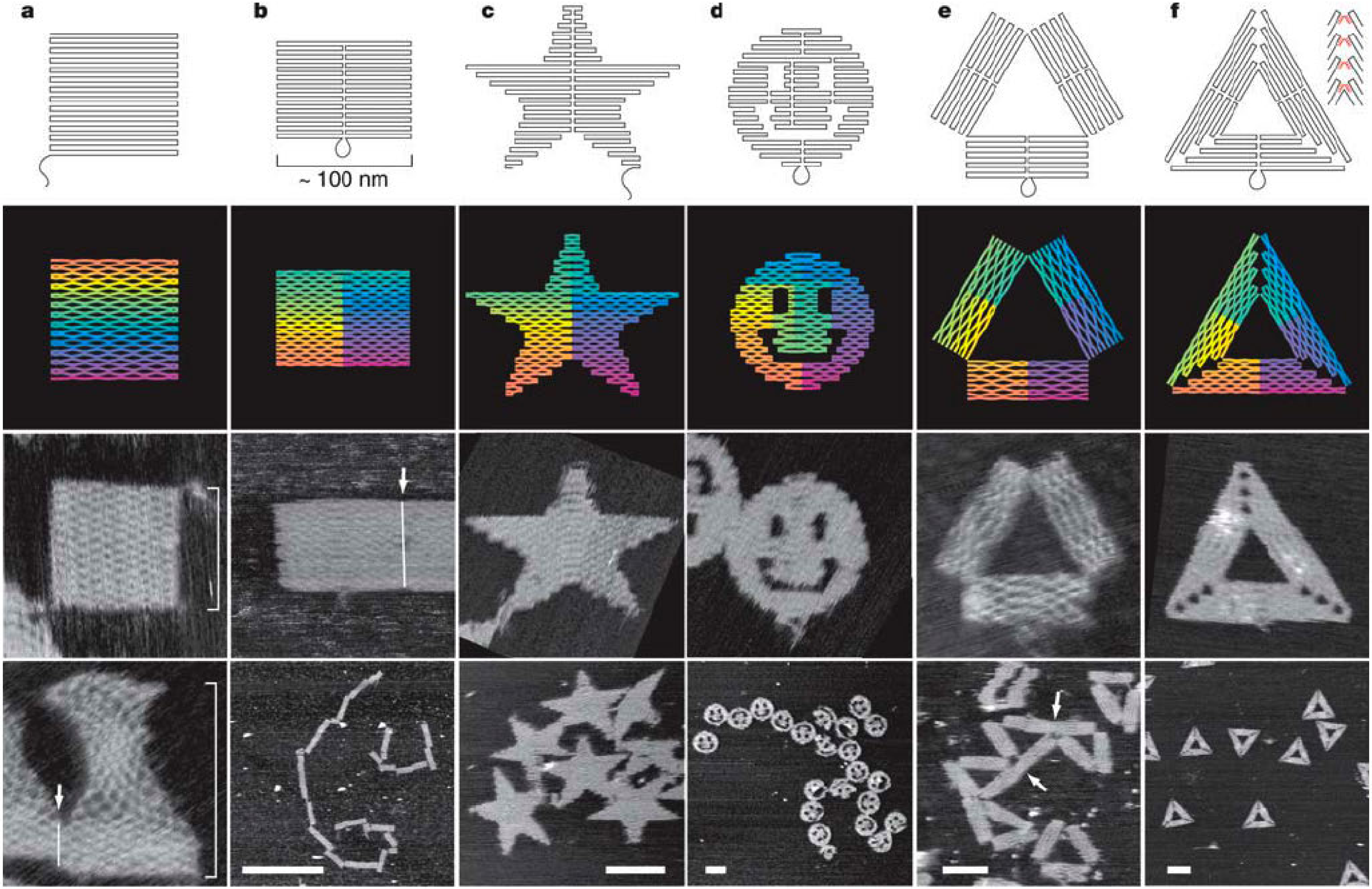
Building a structure with nanoscale precision by **folding** DNA



Viral D old)

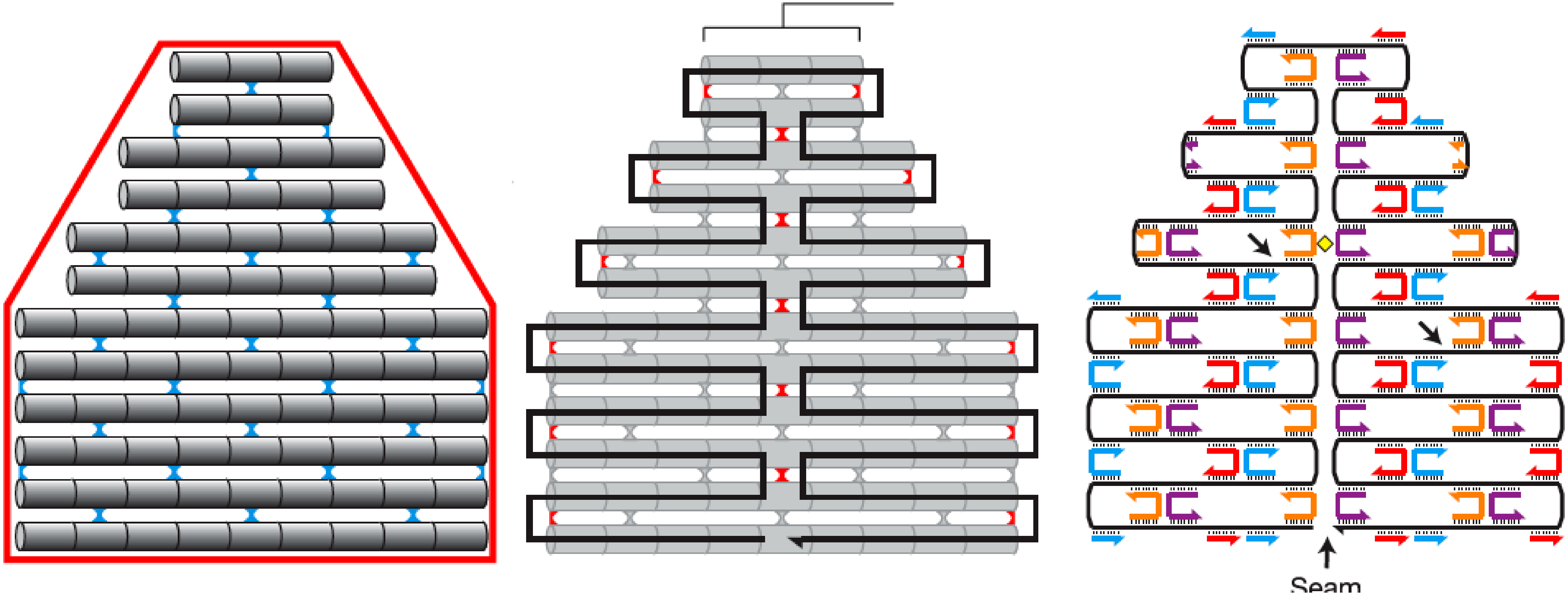
Synthetic (staples)
~30- tides

DNA origami



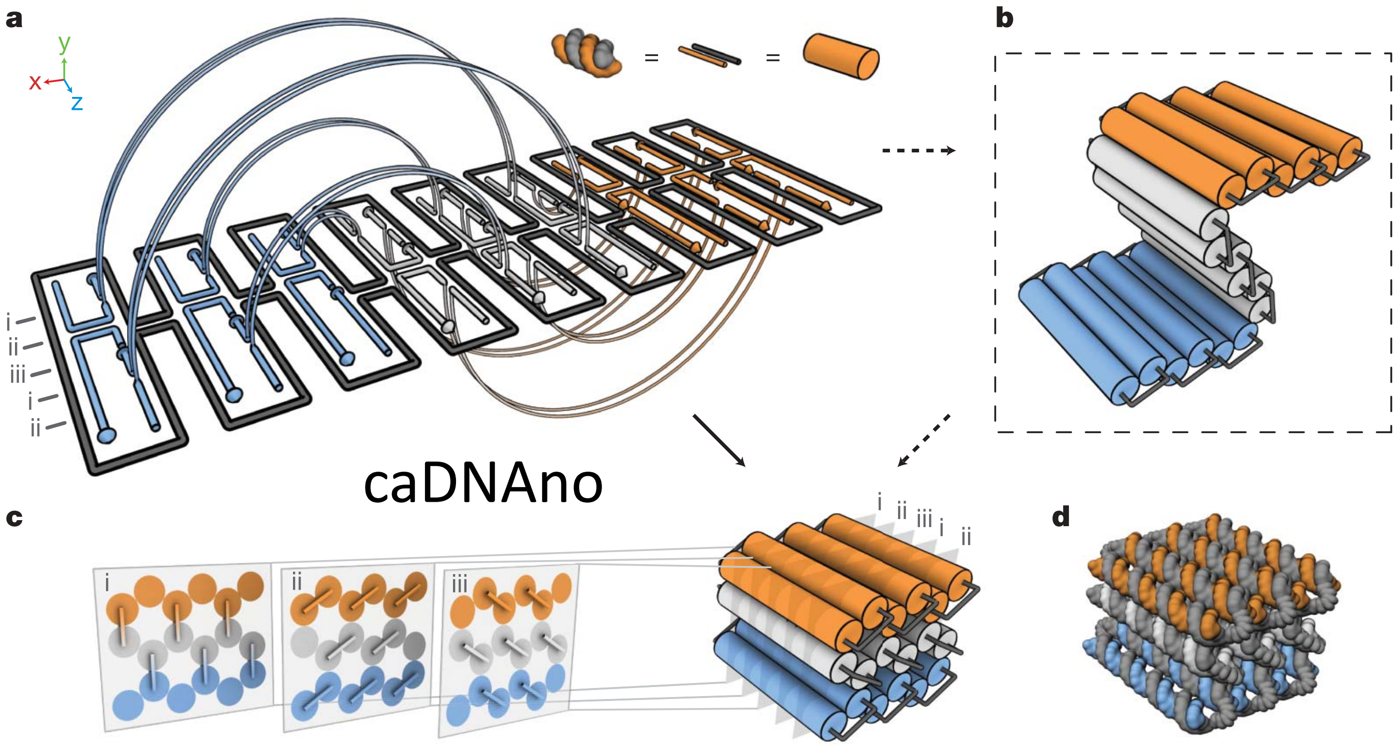
DNA origami

Idea: direct folding of a long single strand of DNA into desired shapes.



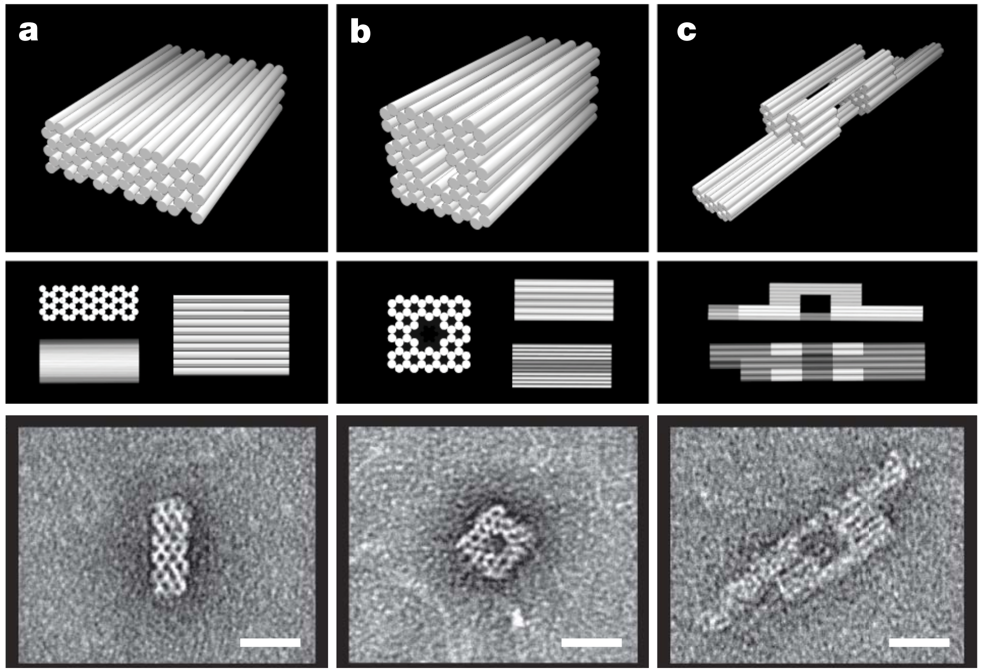
Paul Rothemund (2006), NATURE Vol 440:297

Design and characterization of DNA nanostructures

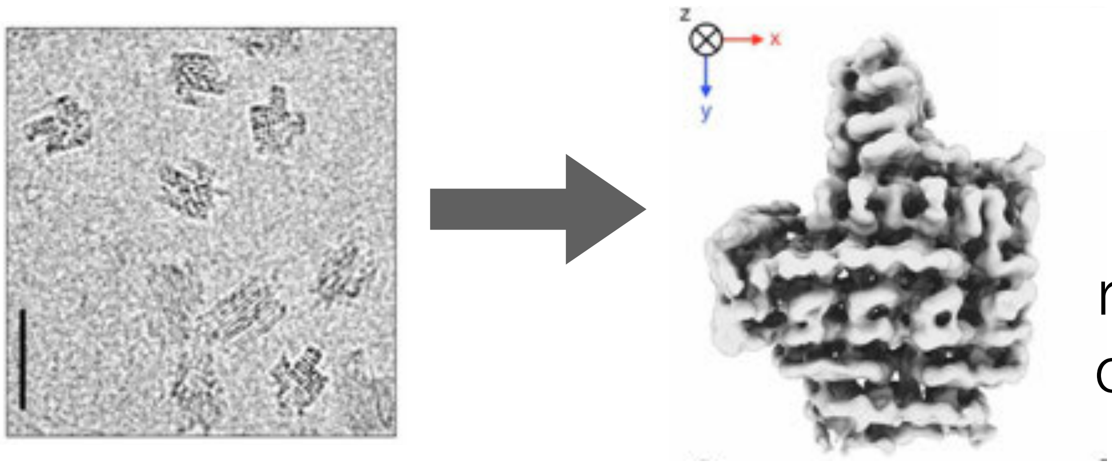


caDNAno

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

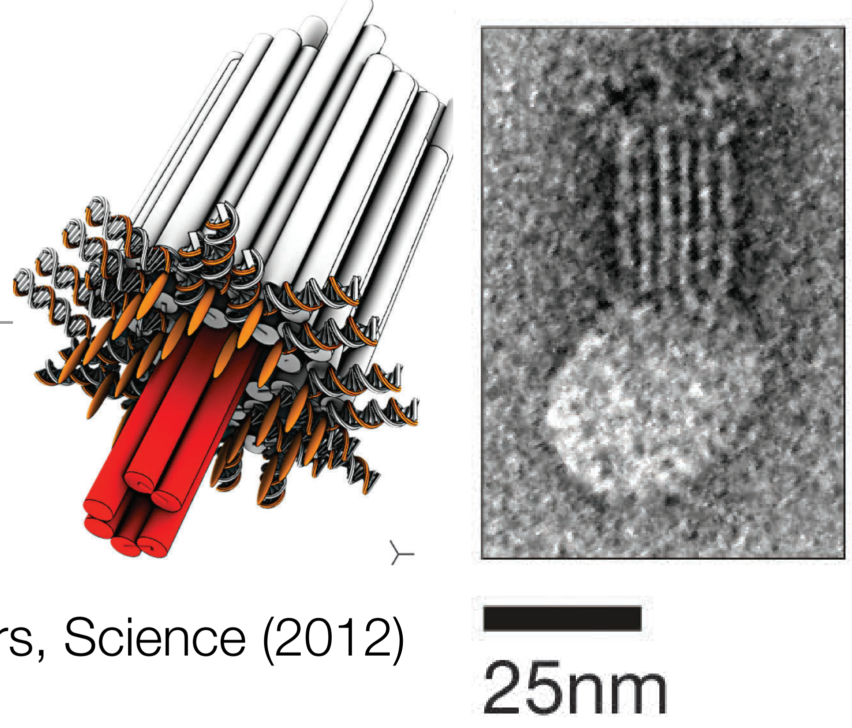


Transmission electron microscopy and/or atomic force microscopy validates the design

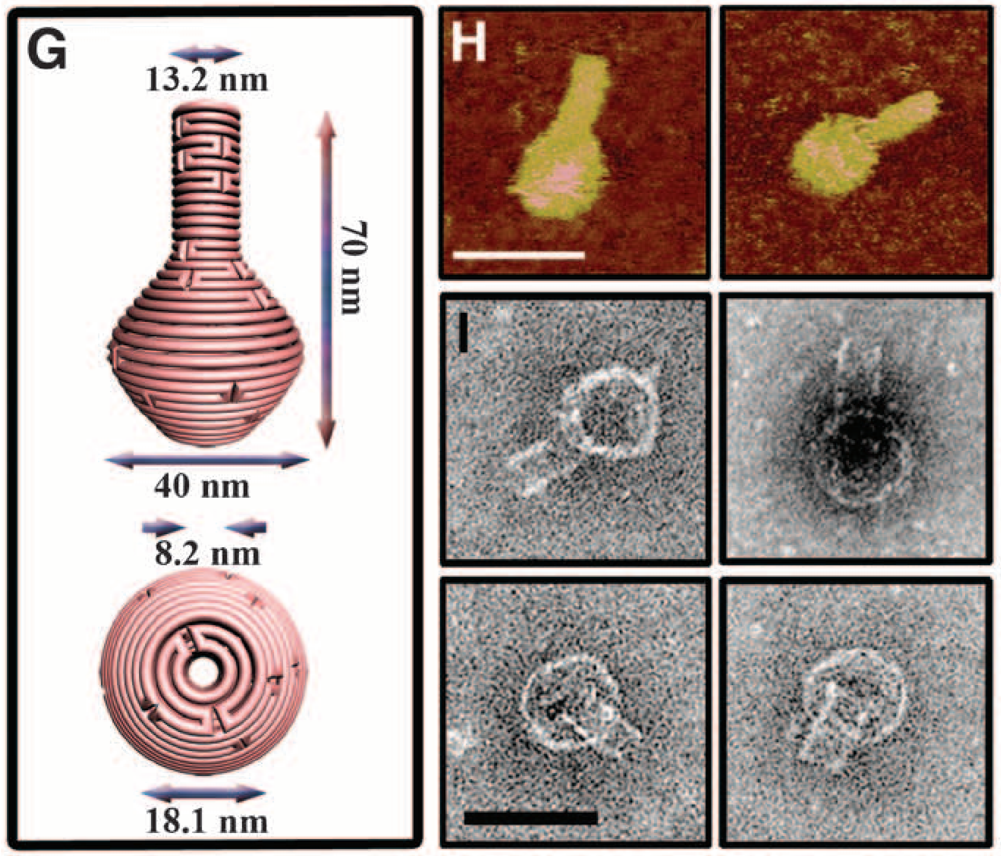


Cryo-EM reconstruction, the only experimentally derived structural model

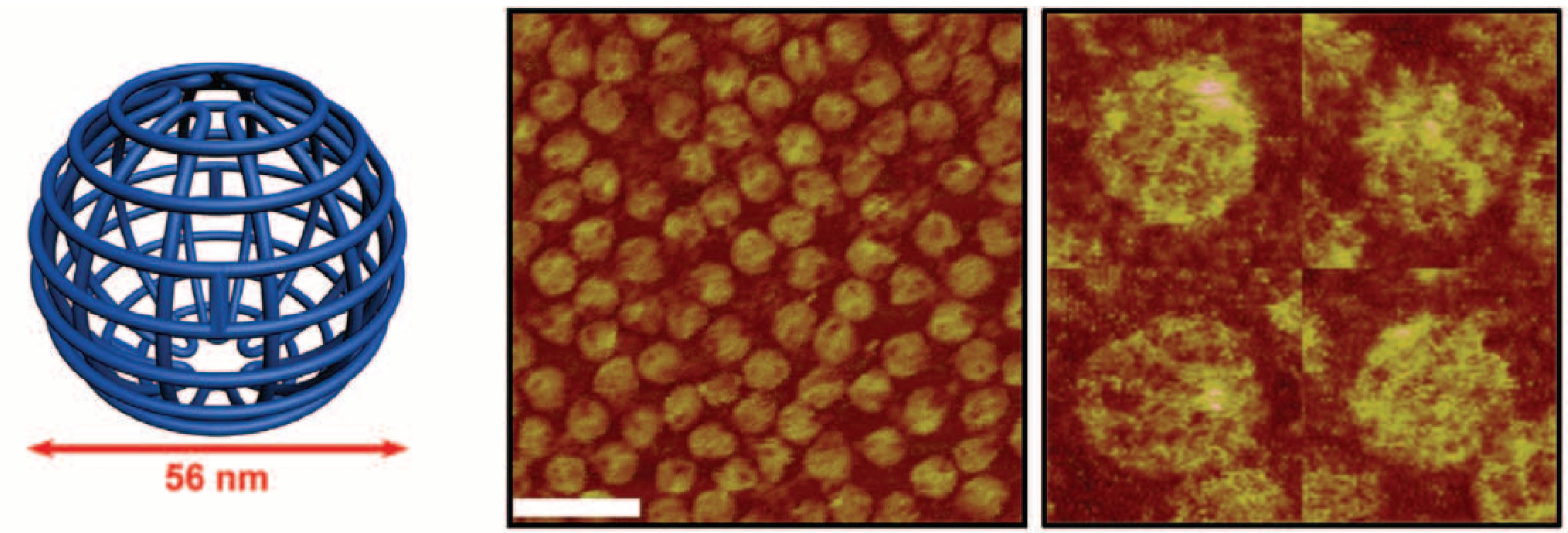
DNA origami structures



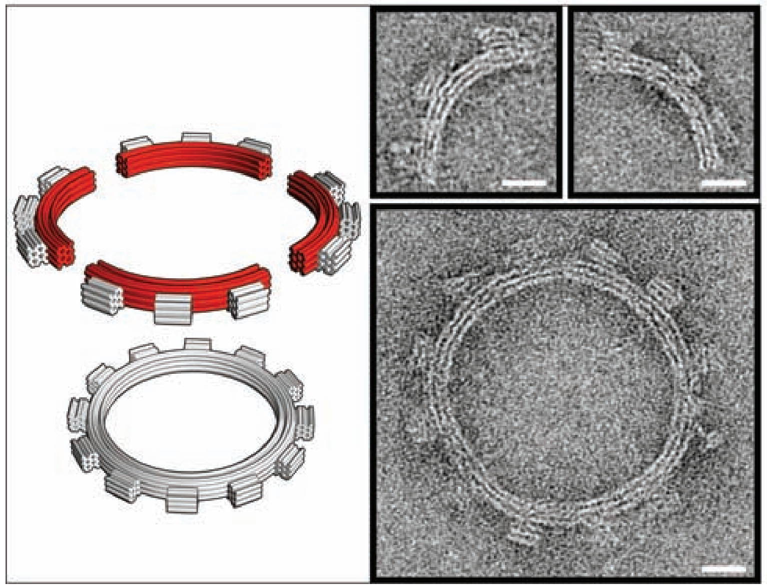
Dietz and coworkers, Science (2012)



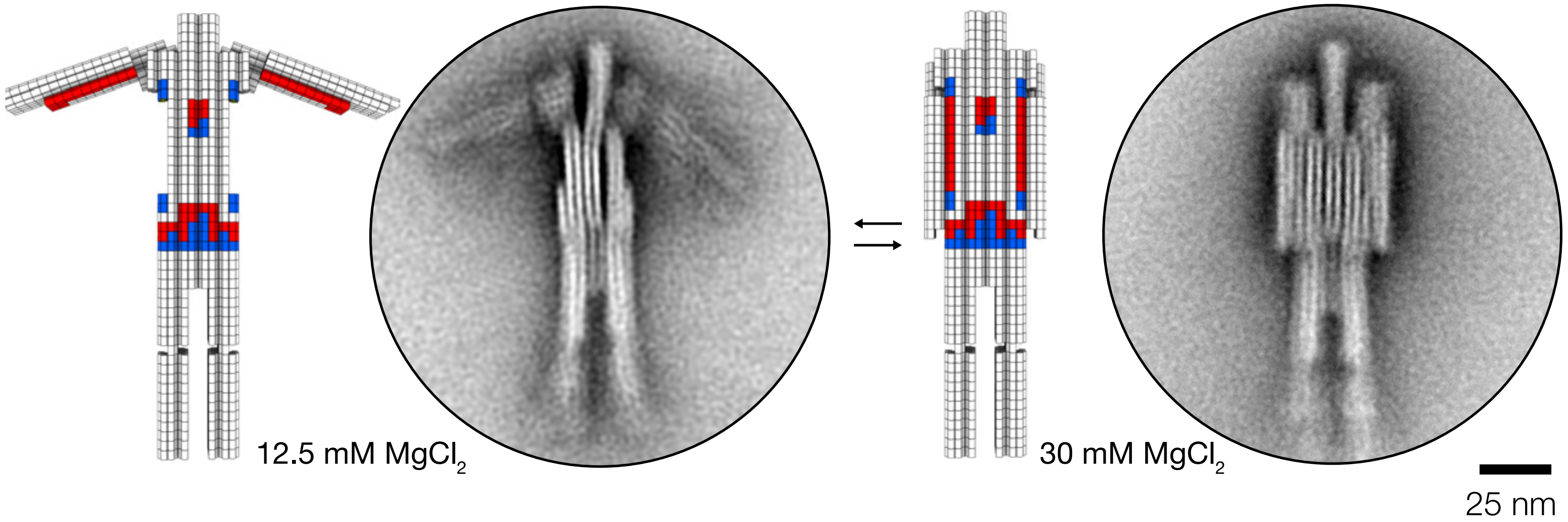
Yan and coworkers, Science (2011)



Yan and coworkers, Science (2013)

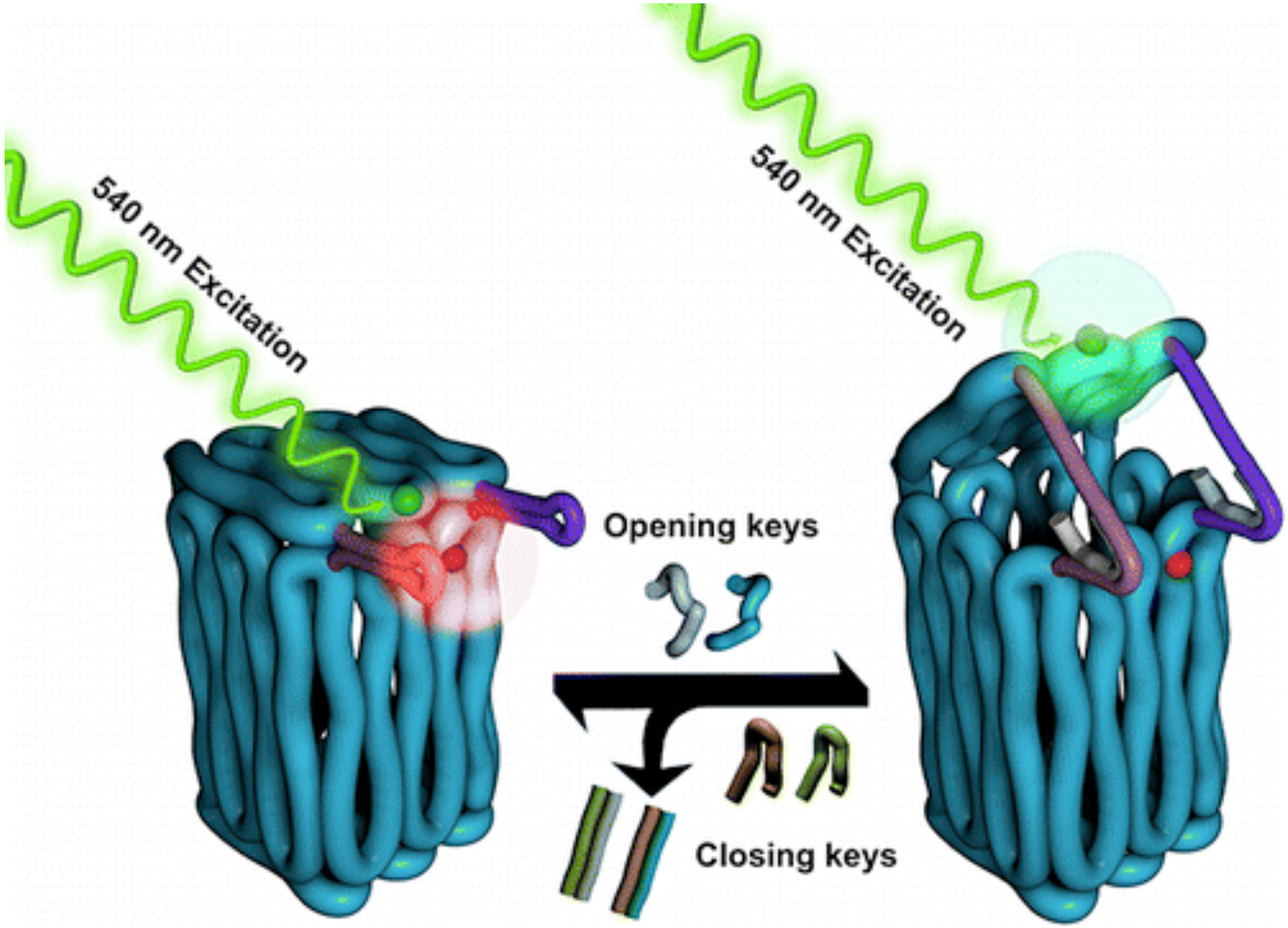


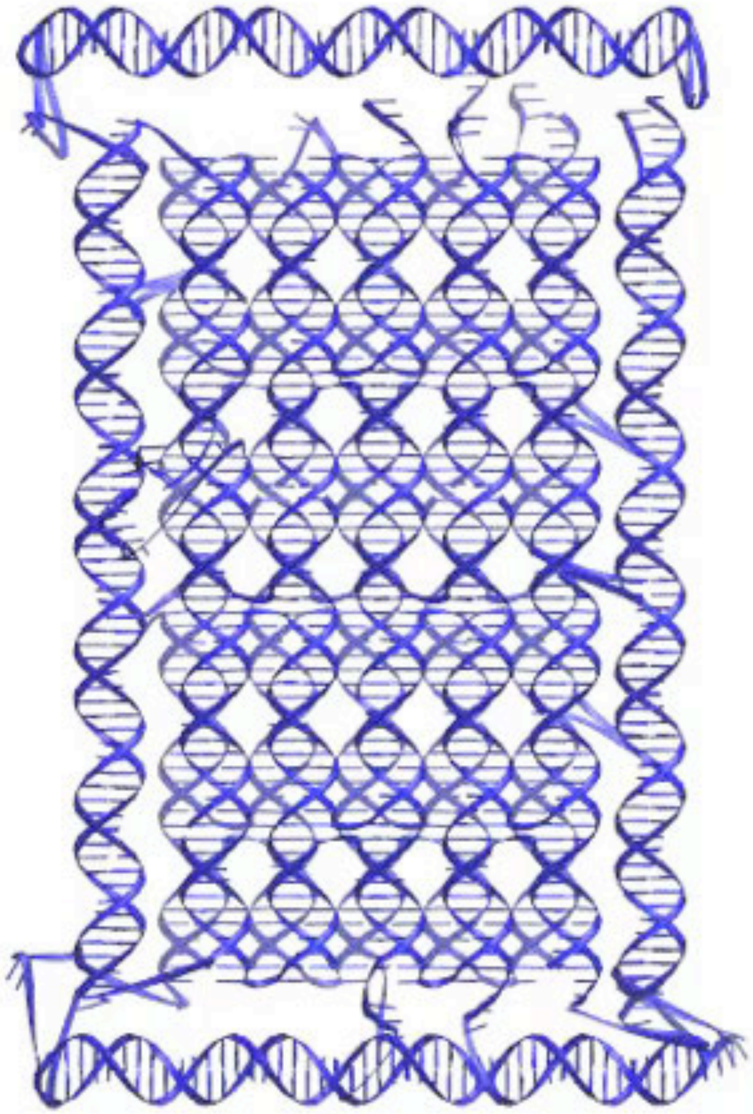
Shih and coworkers, Science (2009)



Dietz and coworkers, Science (2015)

DNA origami box

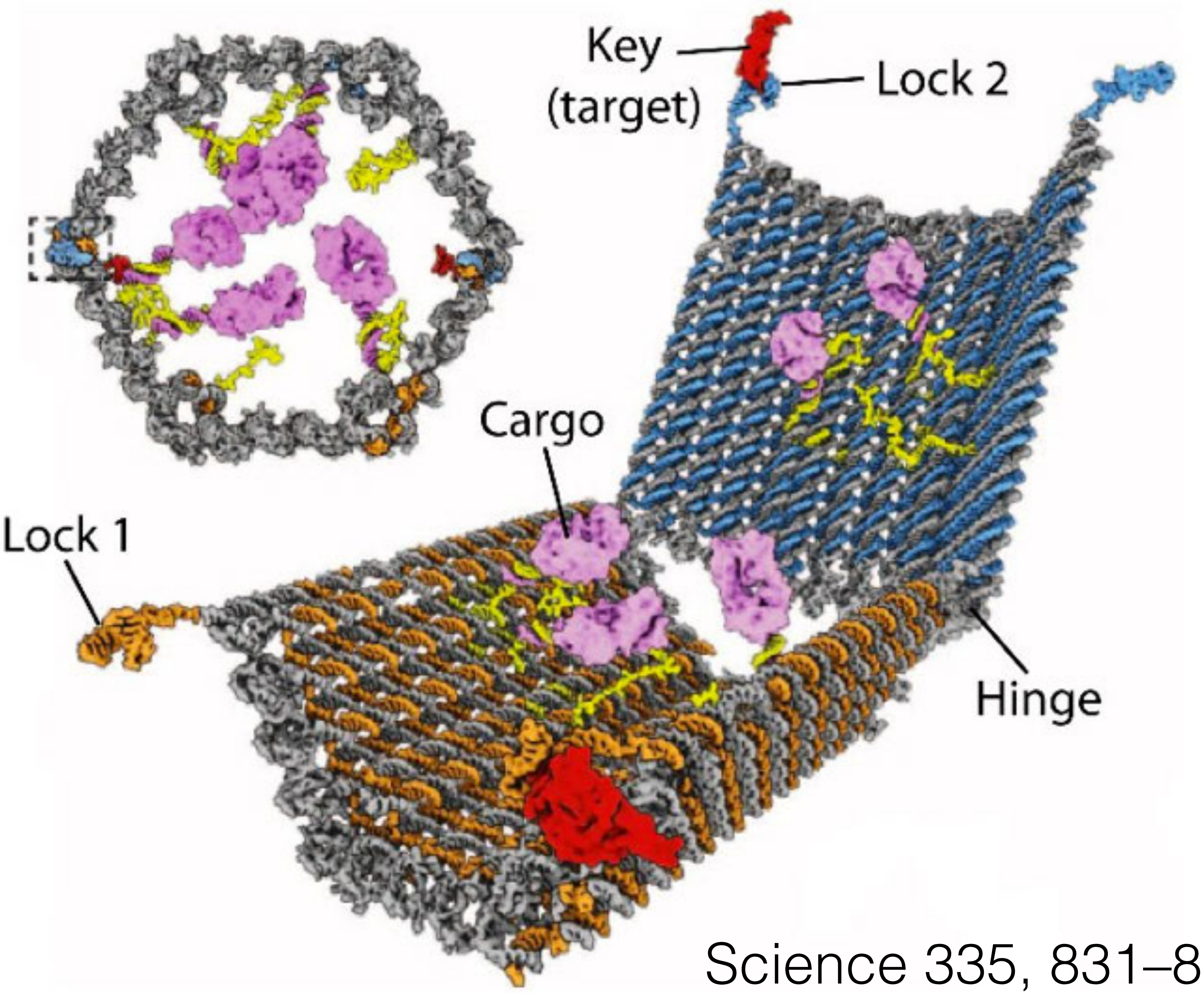




ACS Nano, (2012) DOI: 10.1021/nn303767b

- Designed with Cadnano plugin using the design from Zagdedan et a..
- DNA origami box is 18.0 nm by 13.5 nm by 17.5 nm.
- Solvated system contains ~2.8 million atoms, where ~70,000 are in the DNA origami box.
- The solution contains 10mM Mg ions; gemcitabine (anti-cancer nucleoside), deoxycytidine kinase (protein that activates gemcitabine) and siRNA

Cancer killing DNA robot (2012)

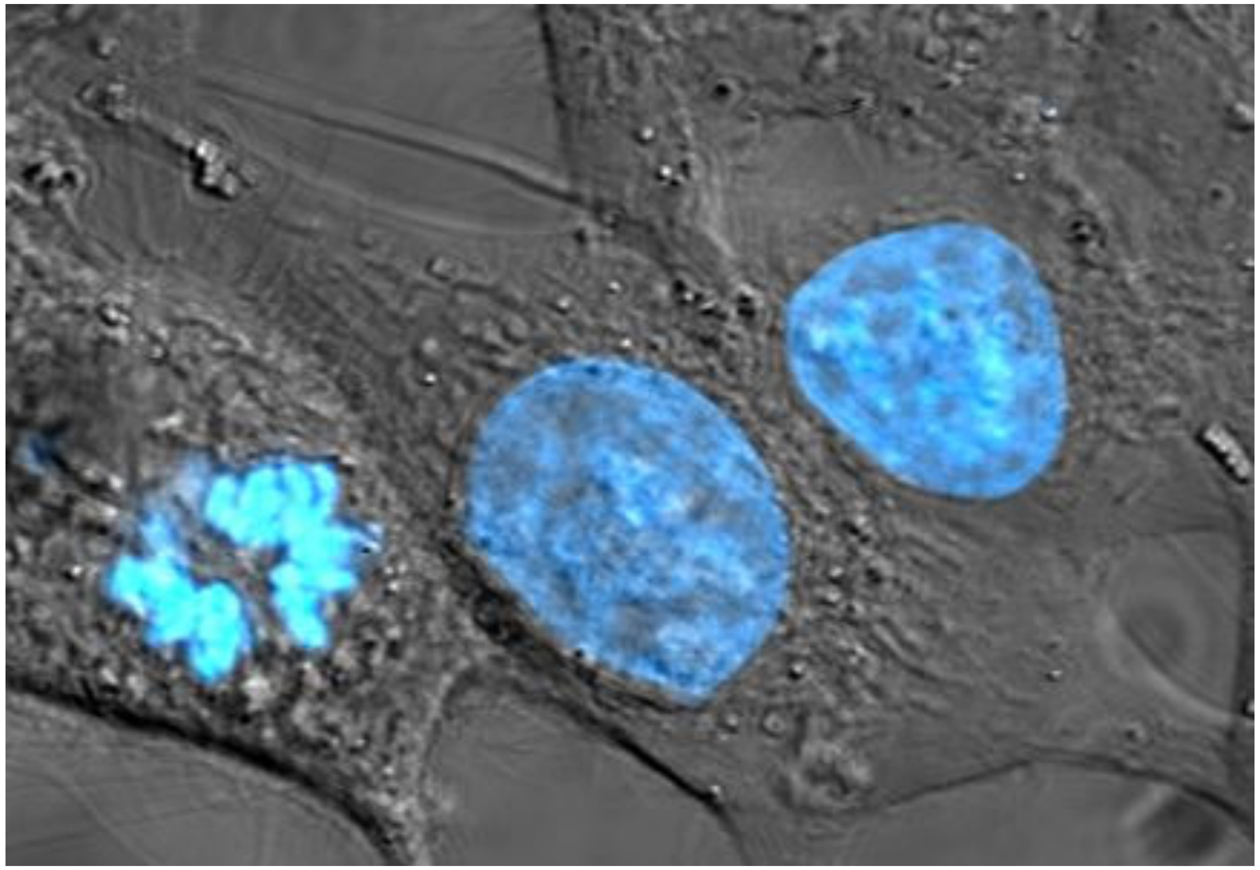


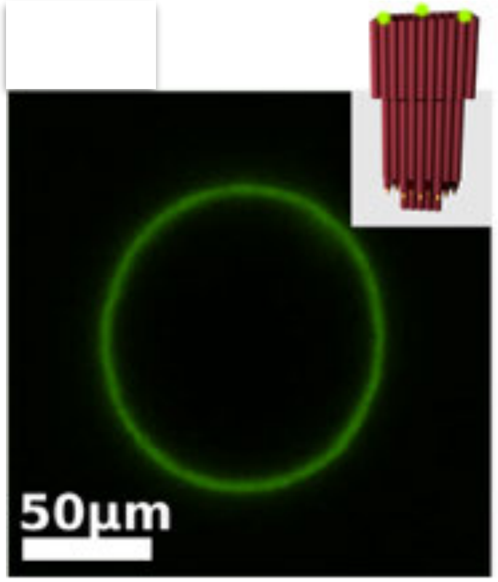
Science 335, 831–834 (2012).

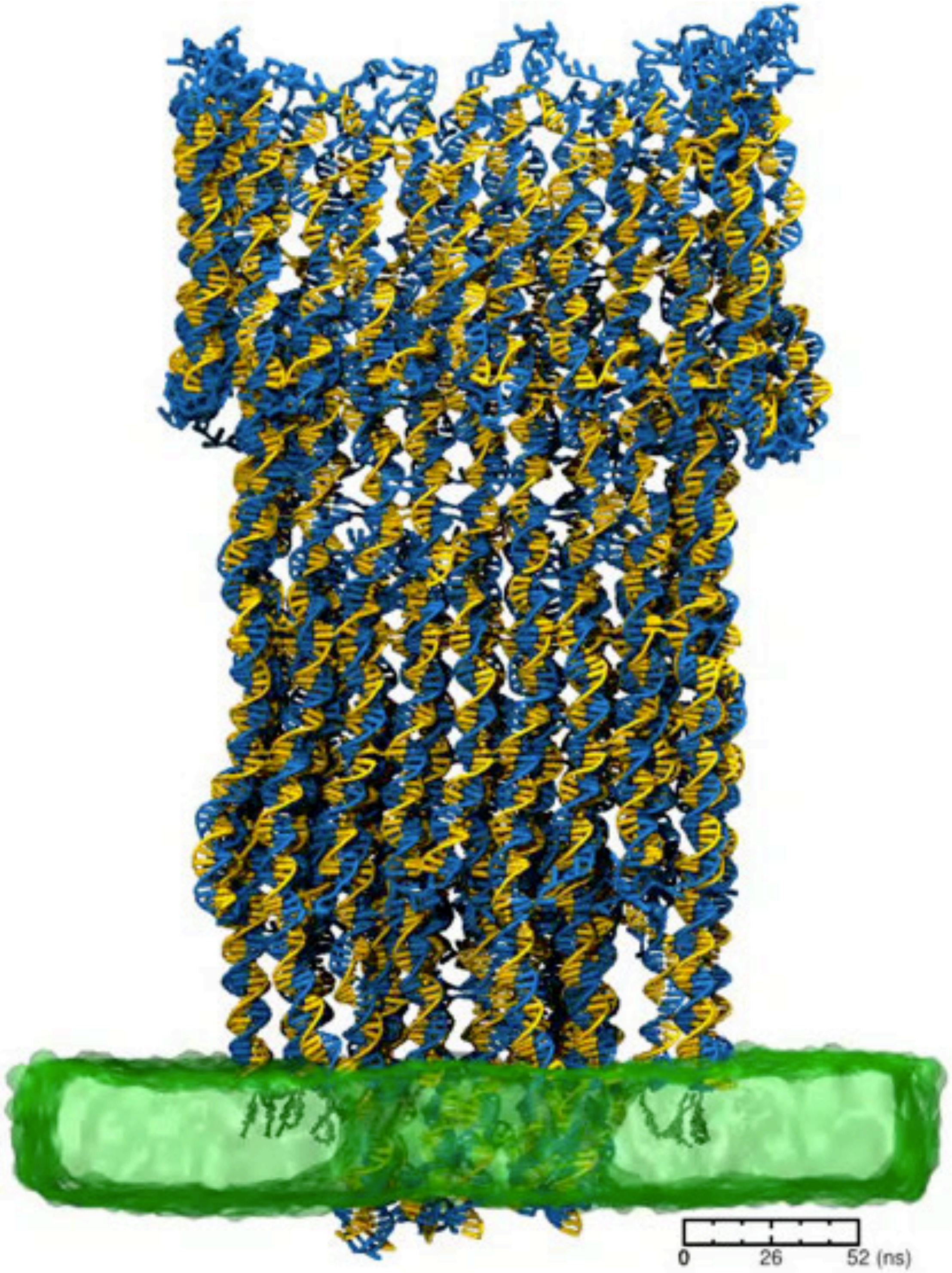
http://www.nature.com/news/dna-robot-could-kill-cancer-cells-1.10047

That's the question from experimentalist.

DNA origami syringe





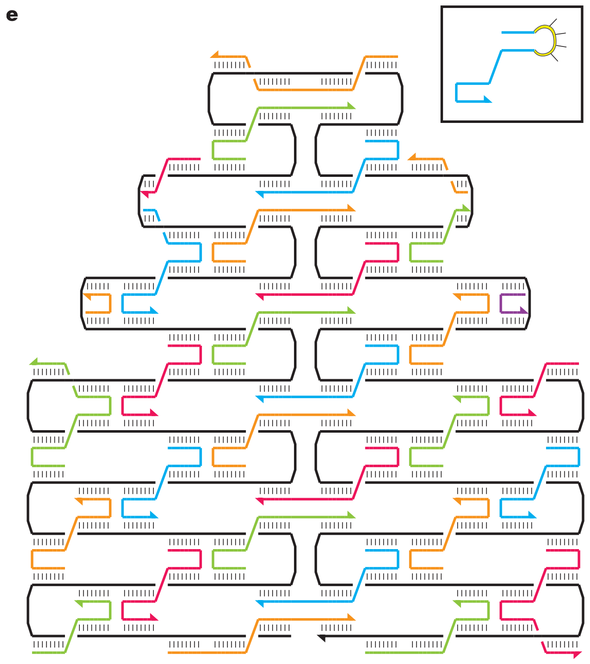


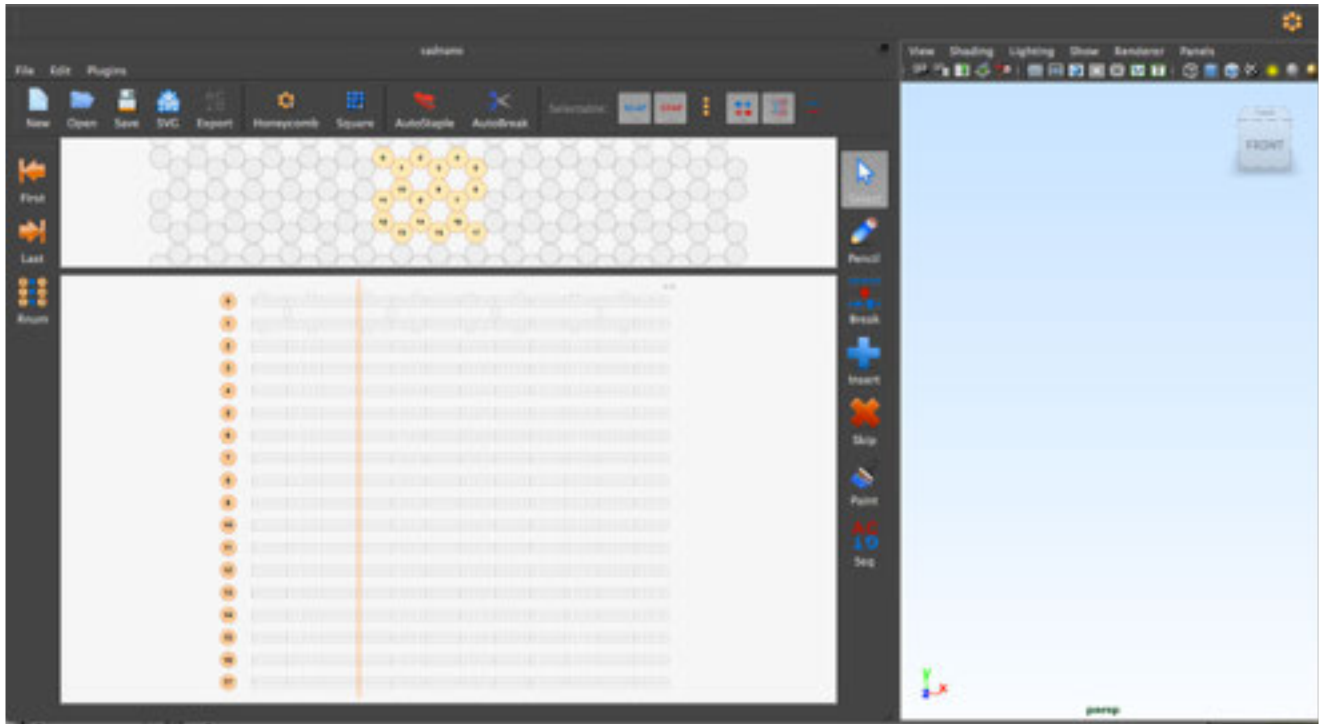
Experiment: Keyser lab (Cambridge, UK)

ACS Nano 10:8207-8214 (2016)

caDNAno

- Computer-Aided Design of DNA origami made by Shih group at Harvard.
- Designed as a plugin of Autodesk MAYA.
- Limited to a design of antiparallel DNA helices in a honeycomb or square lattice.

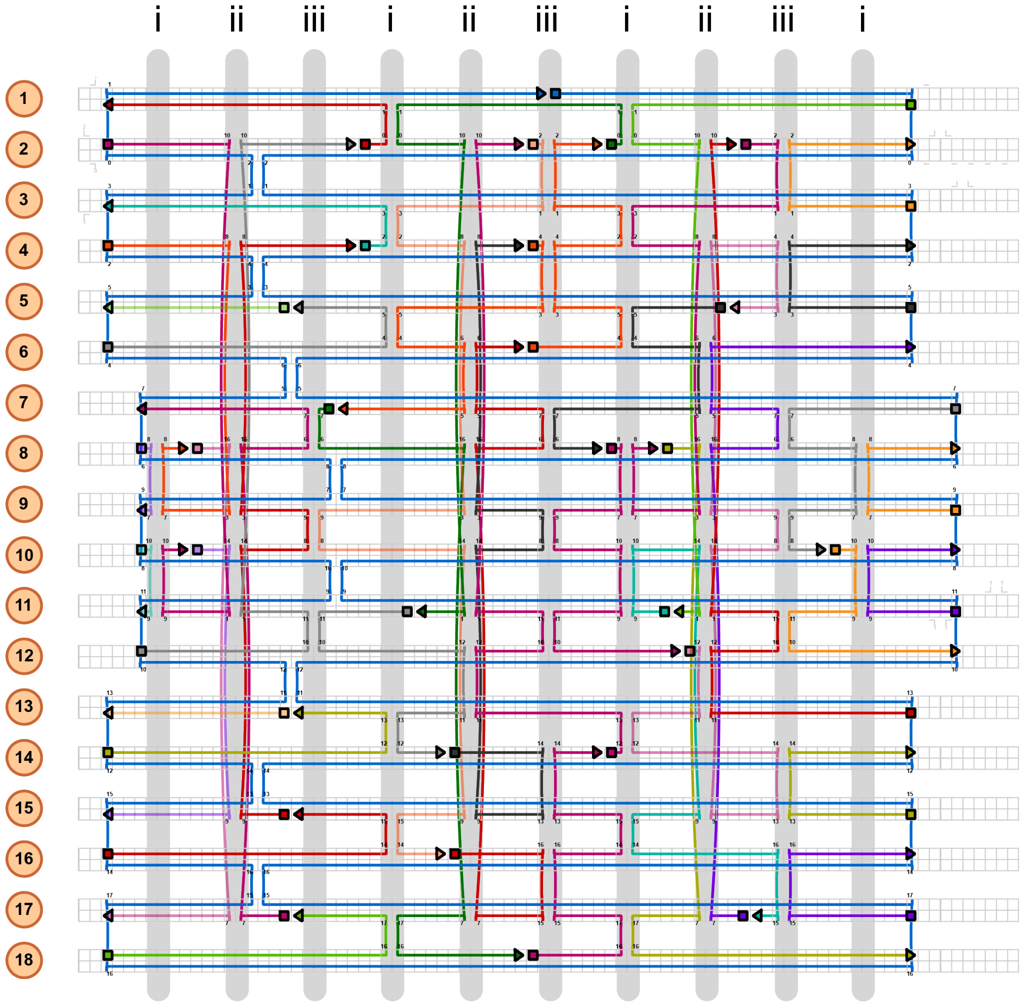


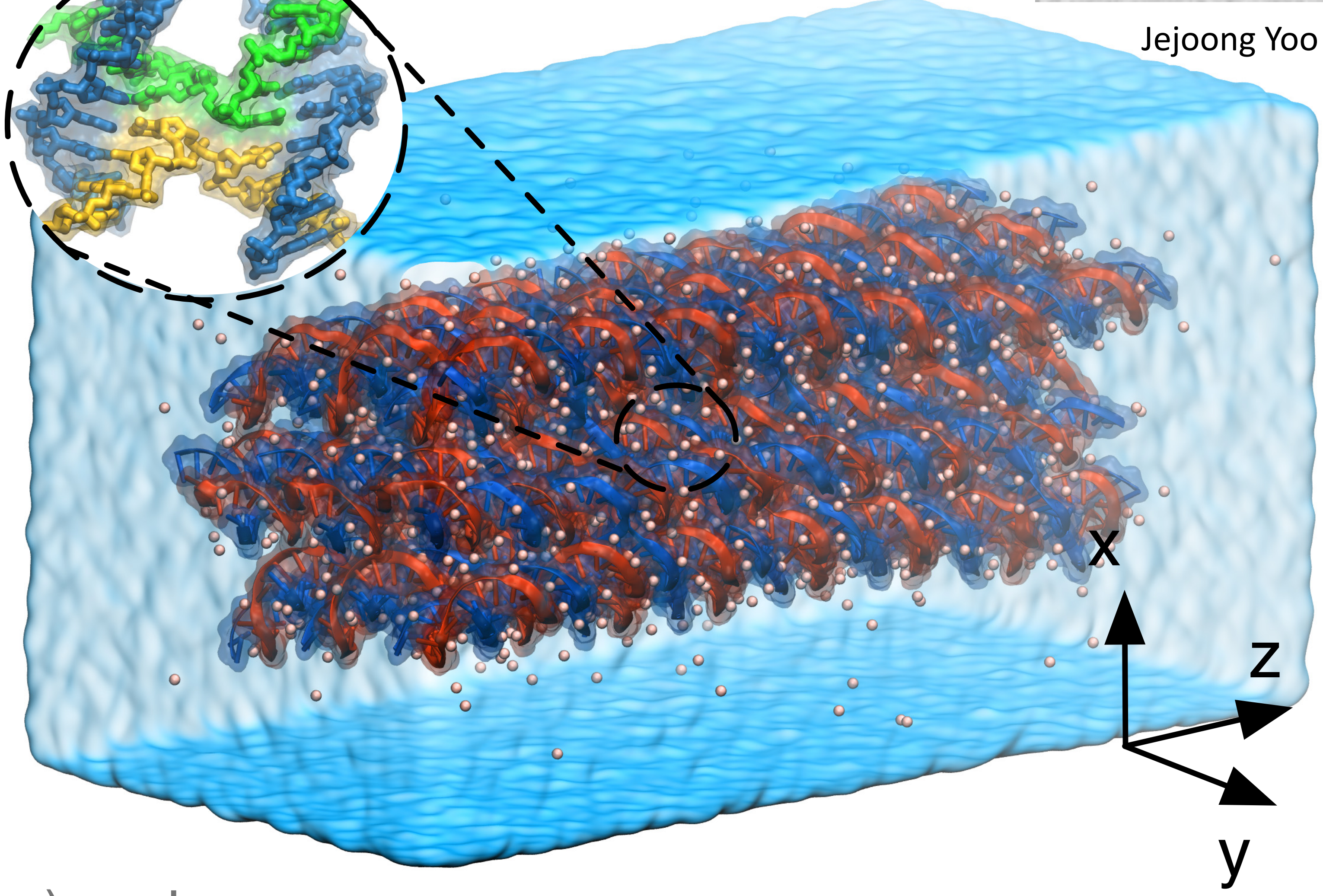


Douglas, S.M., Marblestone, A.H., Teerapittayanon, S., Vazquez, A., Church, G.M. & Shih, W.M., 2009, Rapid prototyping of 3D DNA-origami shapes with caDNAno, *Nucleic acids research*, 37(15), pp. 5001-6.

From caDNAno to all-atom

Jejoong Yoo

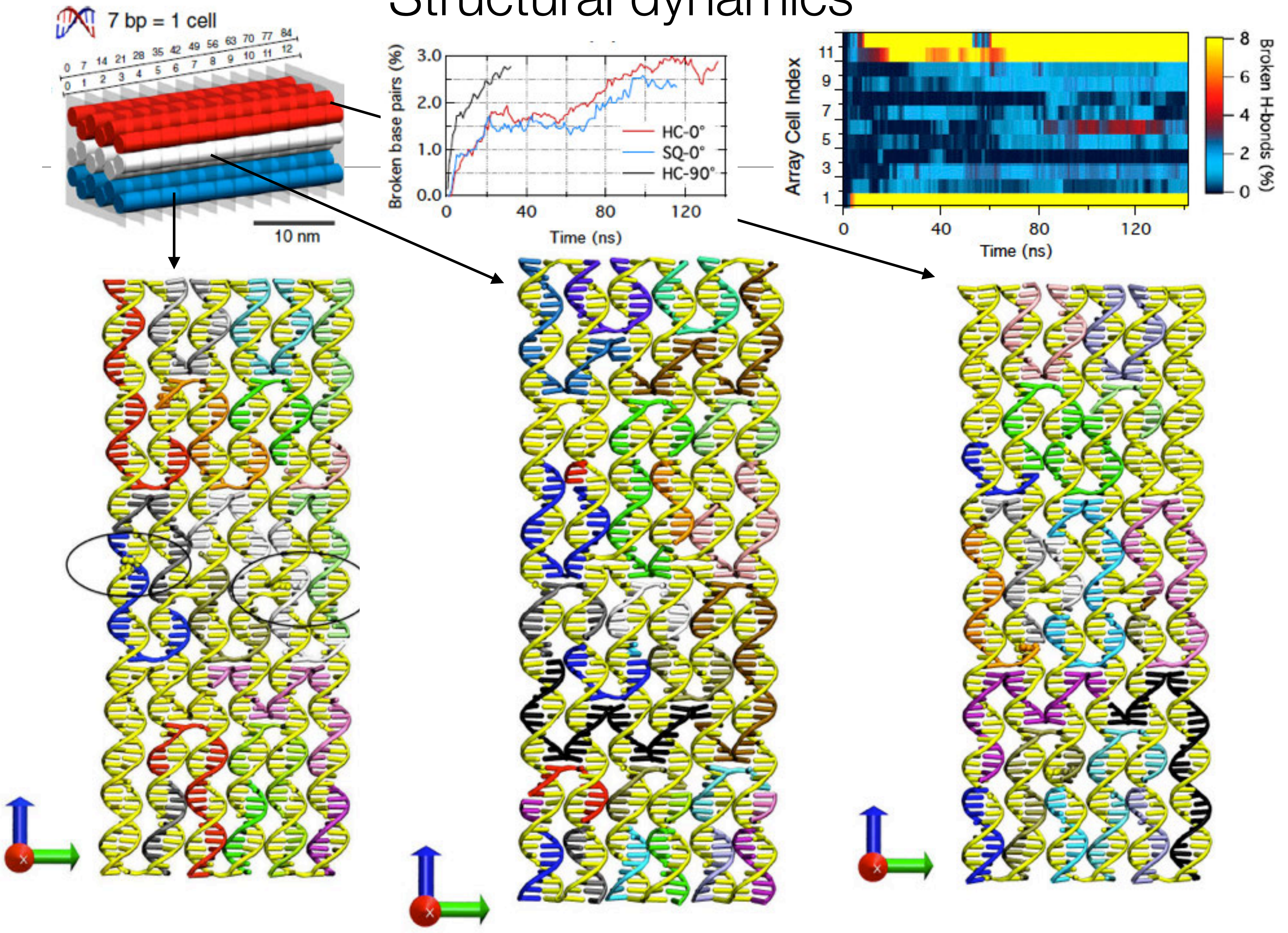




- caDNAno returns topology (json) and sequence (csv) information.
- **cadnano2pdb.pl** combines json and csv files into a PDB file.

✴ CHARMM36 force field
✴ Explicit water
✴ [$MgCl_2$] ~ 10 mM
✴ NAMD
✴ 1 to 3M atoms
✴ 500 to 1,000 CPUs

Structural dynamics

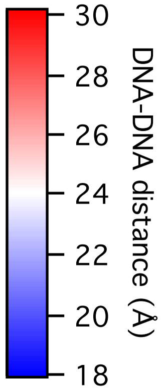


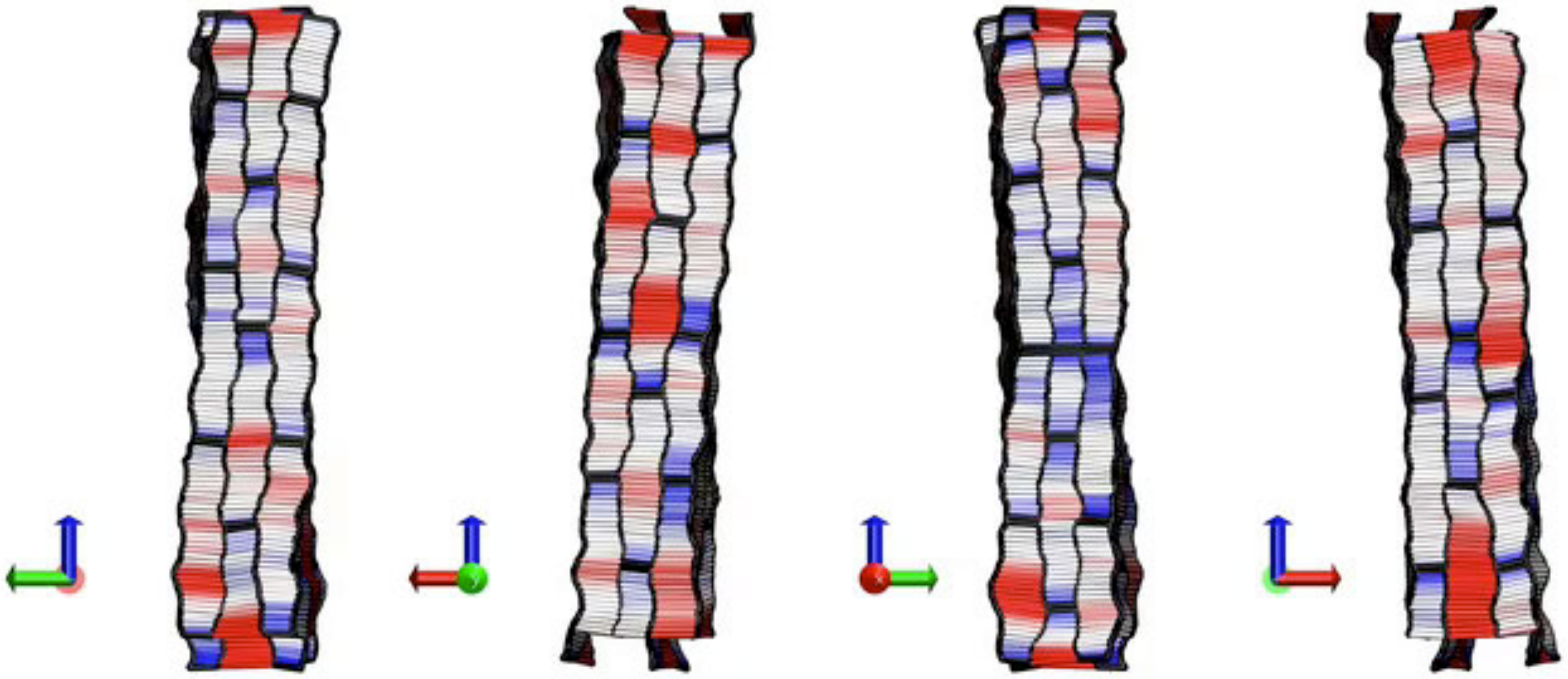
Structural fluctuations reveal local mechanical properties

MD trajectories allow us to compute natural bending and torsion as well as persistence length

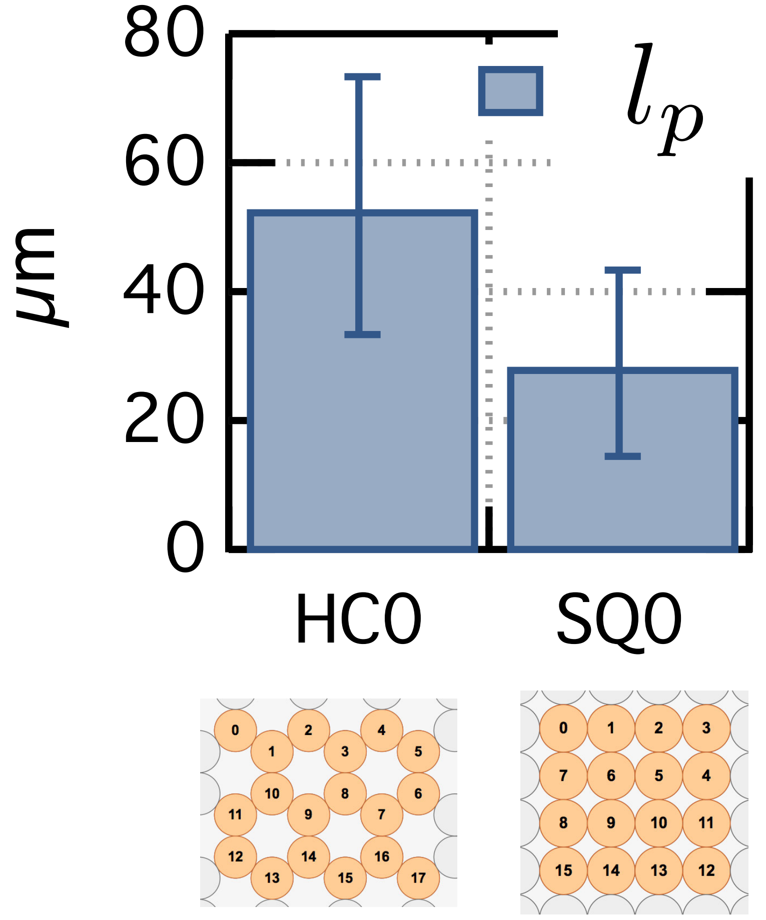
- Inter-DNA distance in color map

- Chicken wire frame represents center line of helices & junction





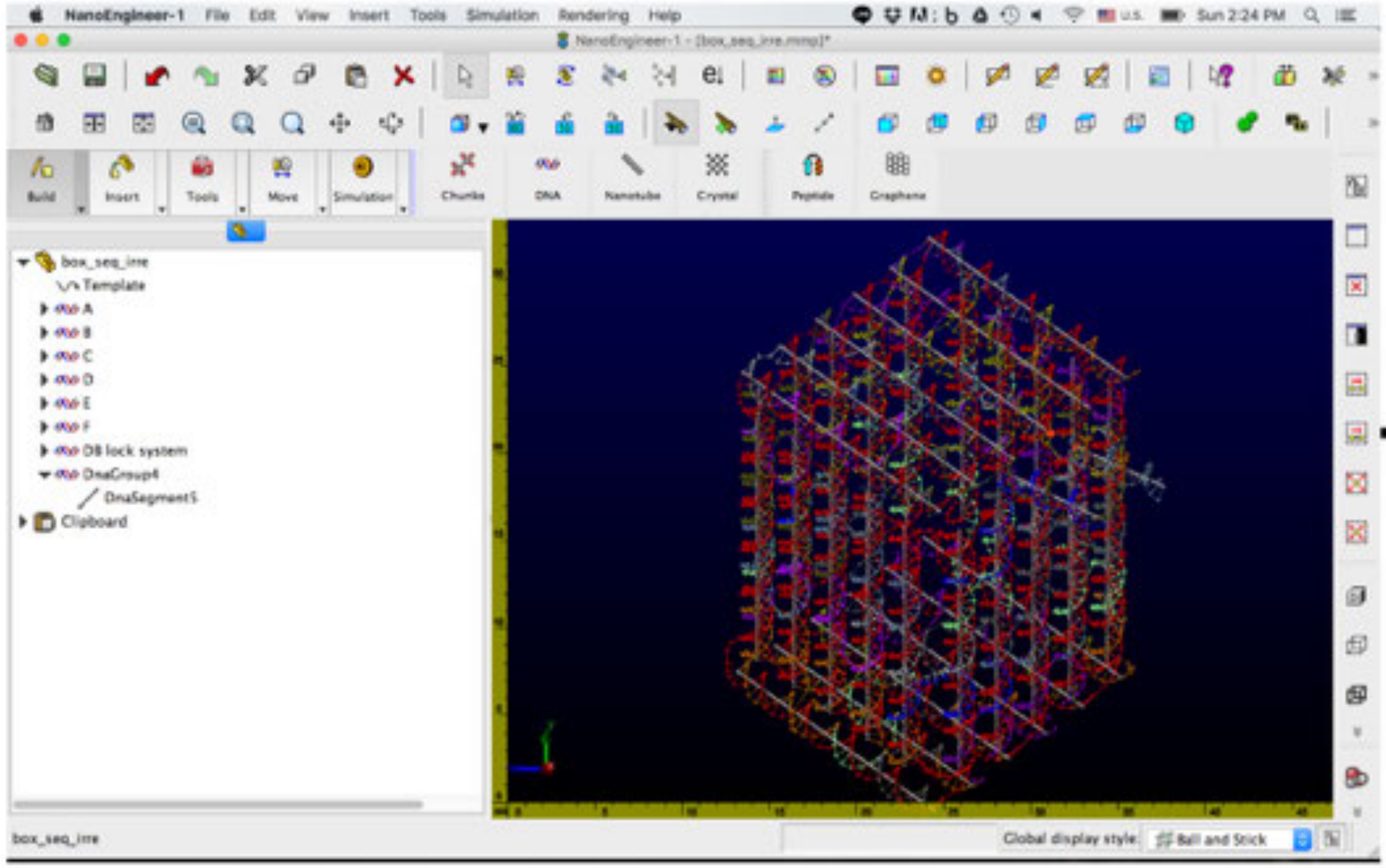
Yoo and AA, *PNAS* 110:20099 (2013)



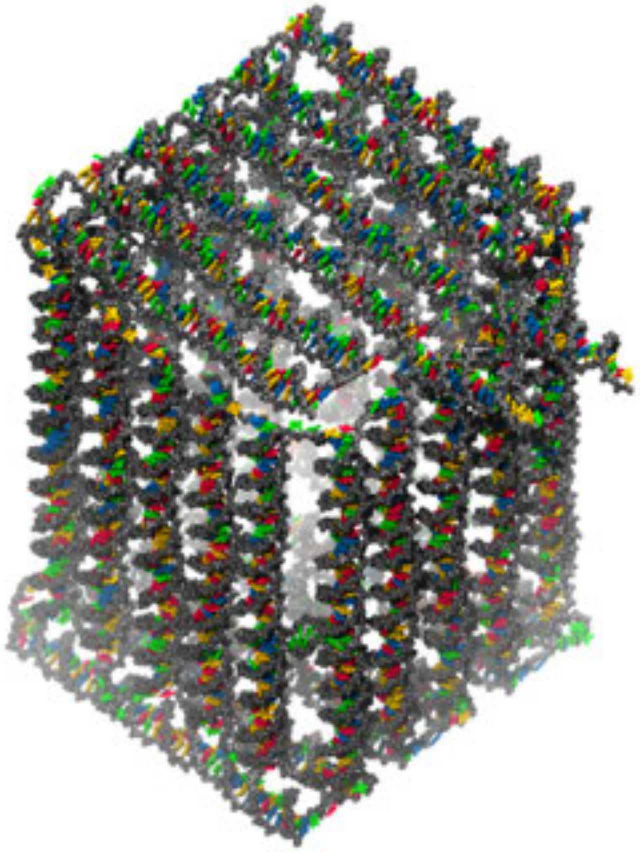
Persistence length of DNA origami

Our simulations predict higher rigidity for honeycomb-lattice design.

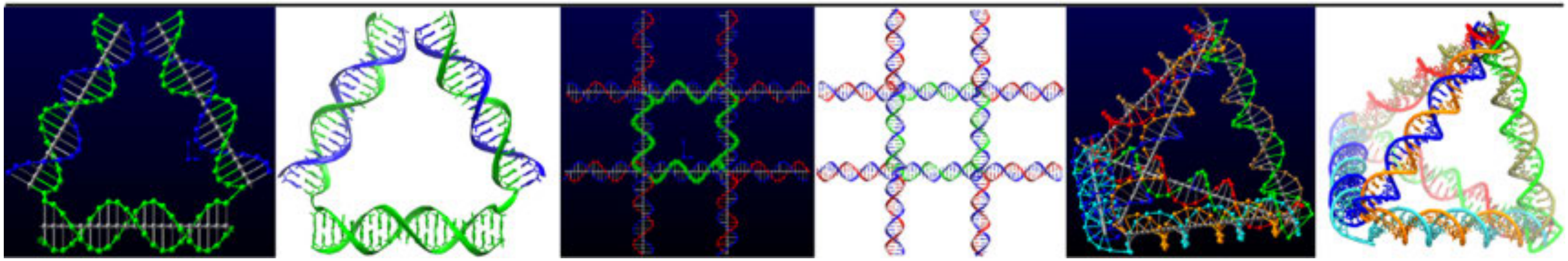
NanoEngineer-1



All-atom PDB

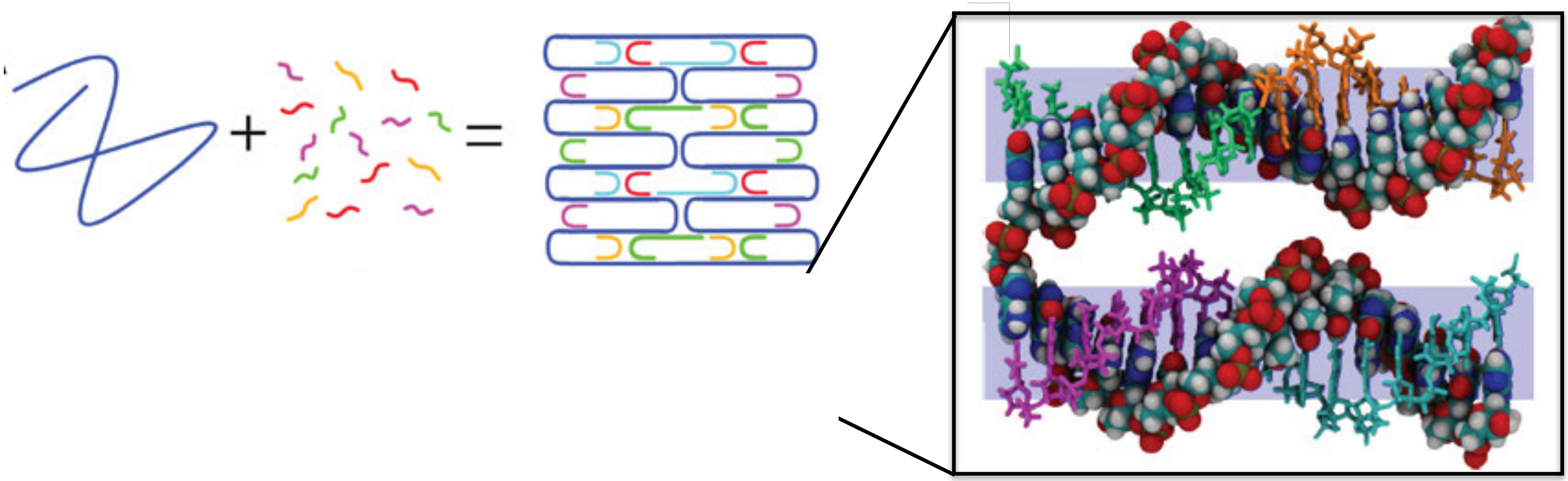


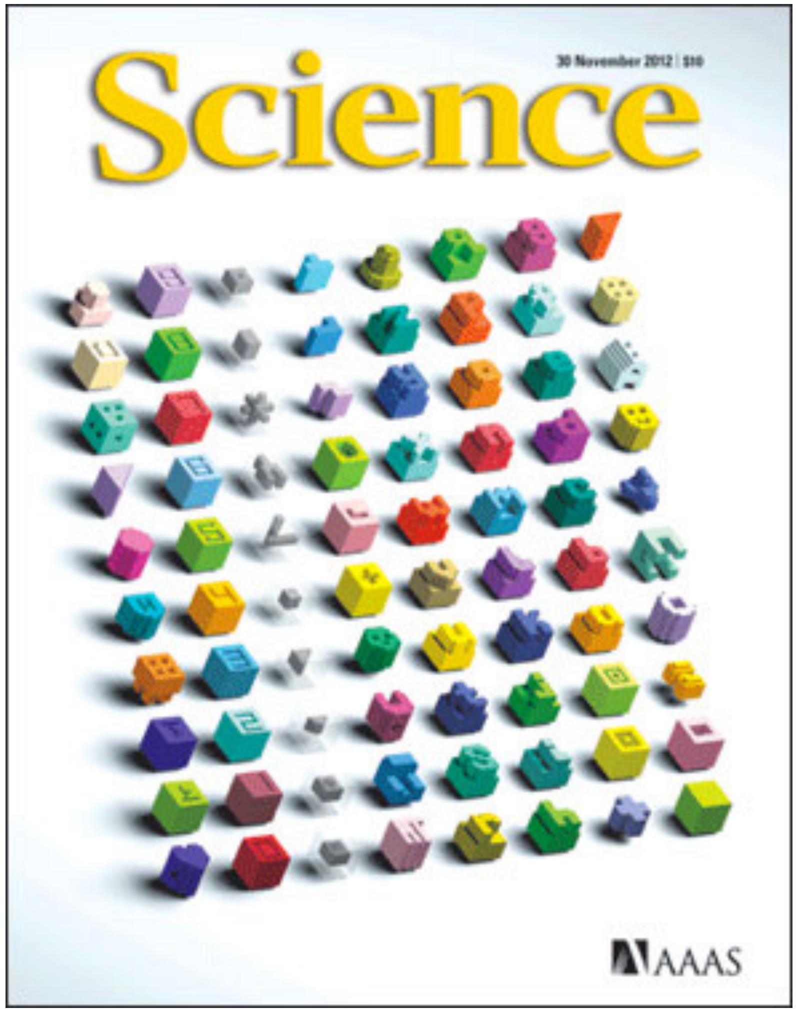
Chen-Yu Li



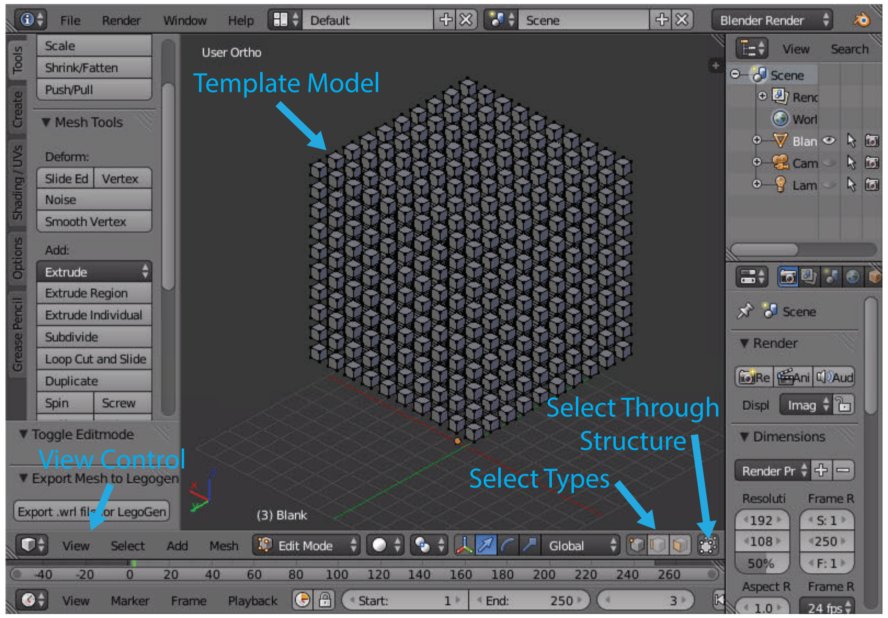
Tiled DNA nanostructures

DNA Origami

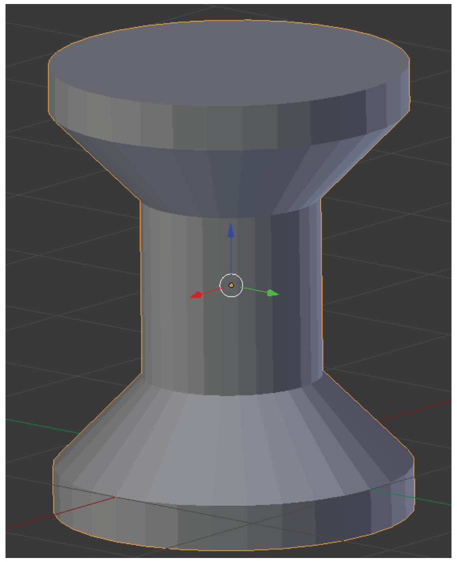




LegoGen workflow @ nanoHub.org

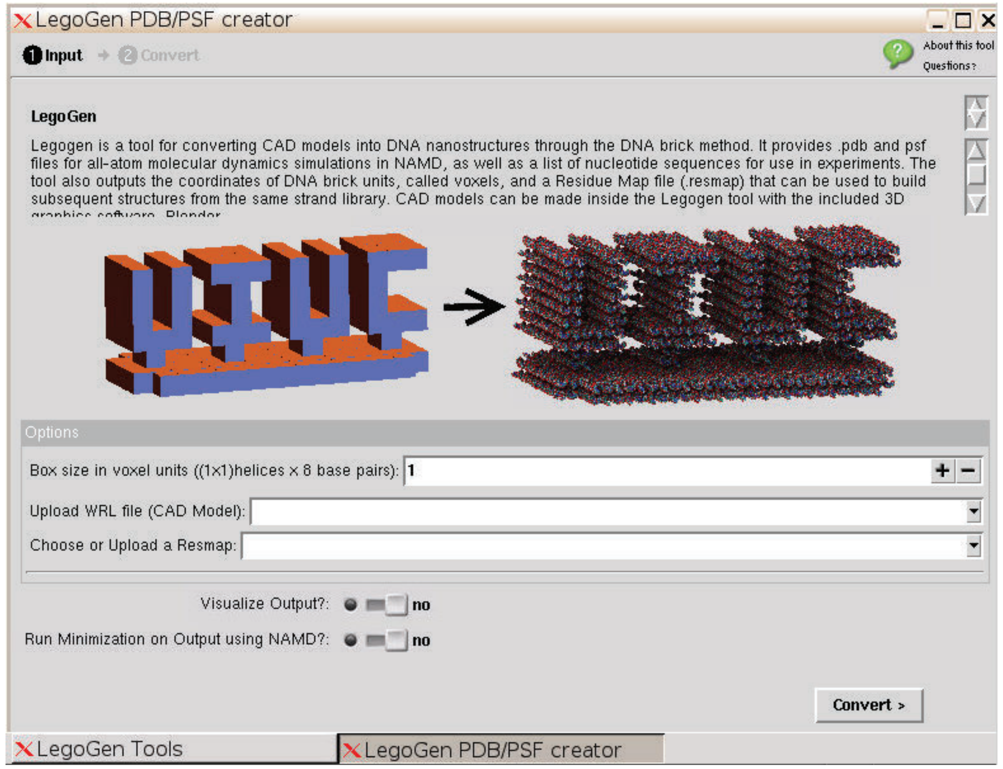


Blender (@nanoHub)



Carved shape

Converter



LegoGen

Legogen is a tool for converting CAD models into DNA nanostructures through the DNA brick method. It provides .pdb and psf files for all-atom molecular dynamics simulations in NAMD, as well as a list of nucleotide sequences for use in experiments. The tool also outputs the coordinates of DNA brick units, called voxels, and a Residue Map file (.resmap) that can be used to build subsequent structures from the same strand library. CAD models can be made inside the Legogen tool with the included 3D graphics software, Blender

Options

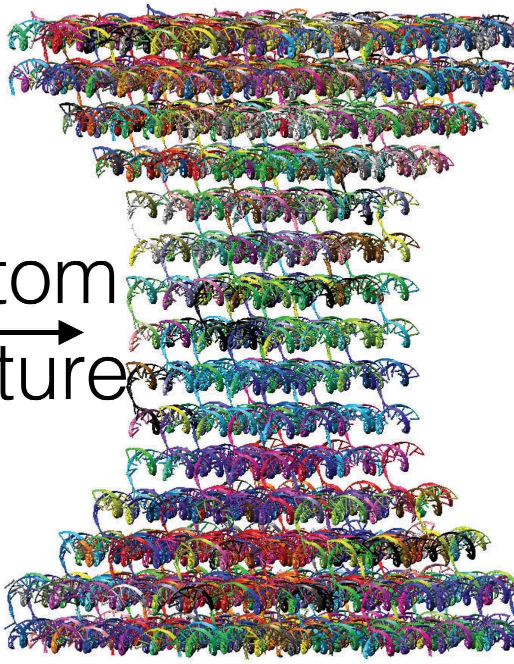
Box size in voxel units ((1x1)helices x 8 base pairs): 1

Upload WRL file (CAD Model):

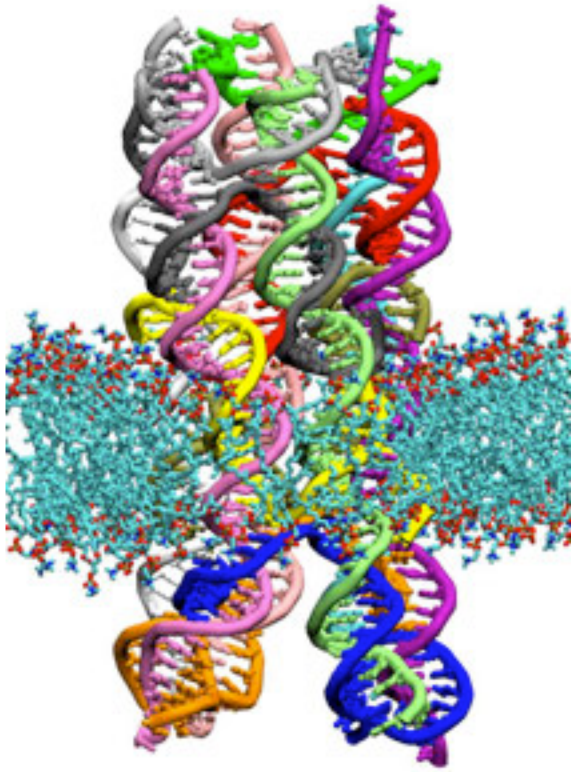
Choose or Upload a Resmap:

Visualize Output?: no

Run Minimization on Output using NAMD?: no



All-atom structure

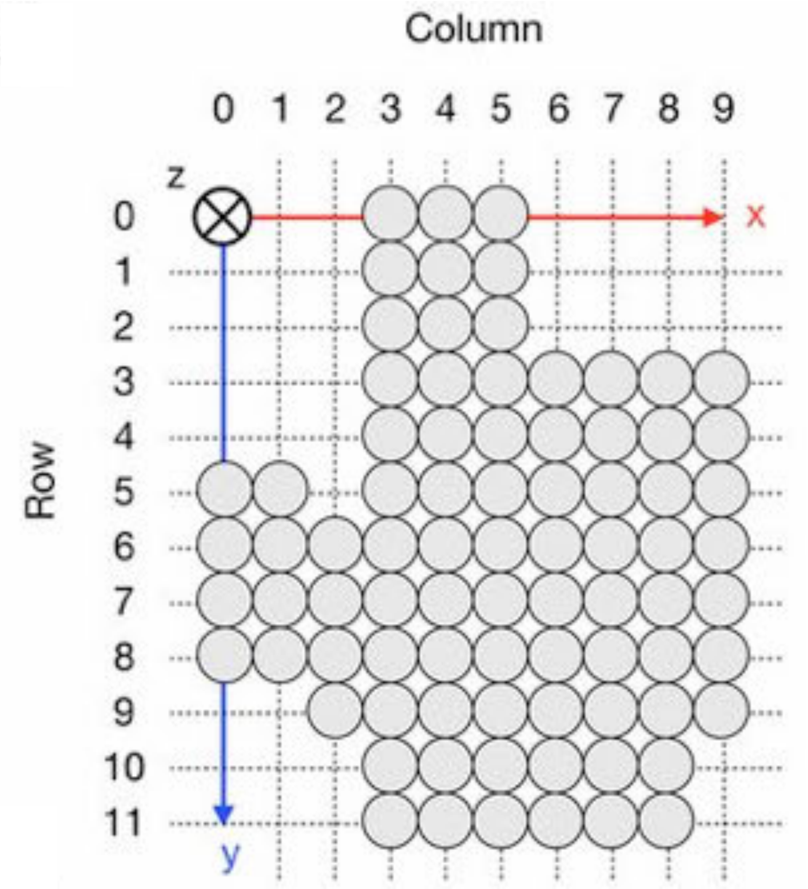


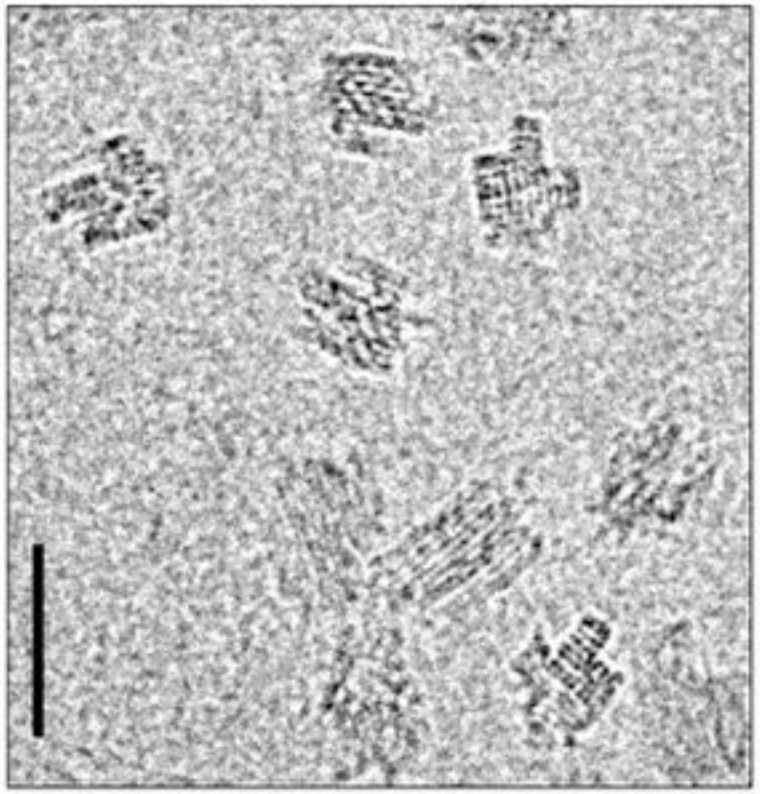
MD simulation

1 ATCAATAGGCG
2 AATGGGCGAAAAACCGTCTGGACTCCAGTT
3 ATCGTCATAAATATTCCGTGCCAGCAGGGTGGAGG
4 AACAGTTCCTCACTGCGCAACAGCAACGACGGCGC
5 TCAAAAATCATGAGTGAGATTAGCAAACGCCAGGGCT
6 TATCGCGTTTTTGCGGAGAGAGTTACATTTCGACG
7 GCGGATCCCTGACTATTATAGT
8 AGTAAGCAAACTTGGGCGCCTGCATTACTGTGTGA
9 TTTAAGCCCCAGCCAAAAGAACAGCCAGCTTTCC
10 AATTGCTGAAAGAGGAAGGGCAAAGACTAAC
11 GAGTGTTGTTCCAGTTTGGAATTATAAATAGGCG
12 TTAGTAGGTTTGATAAGAGGTCATTTTAATTCTGA
13 GGAAGTTTCATTCCATATAATGTTTAGCAAT
14 GCAACTAATGAAAAGGTGGCA
15 CCGGAAACCAG

Experimental realization

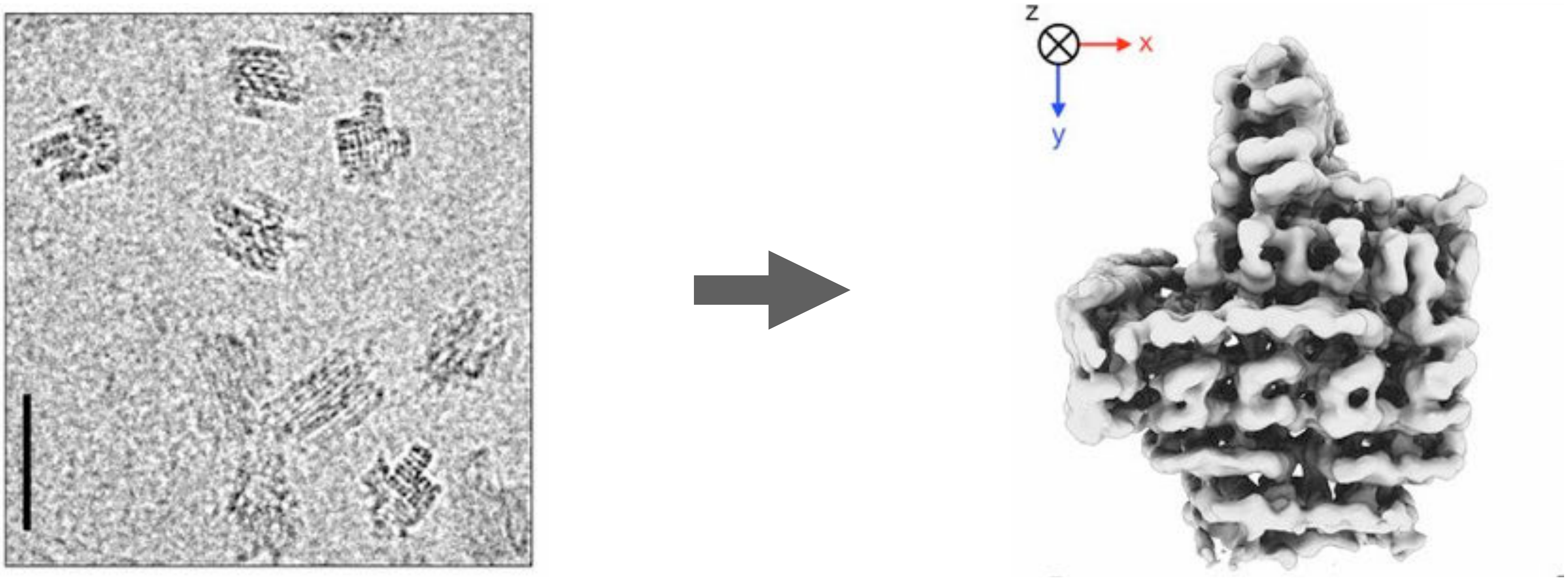
Cryo-EM reconstruction versus all-atom simulation





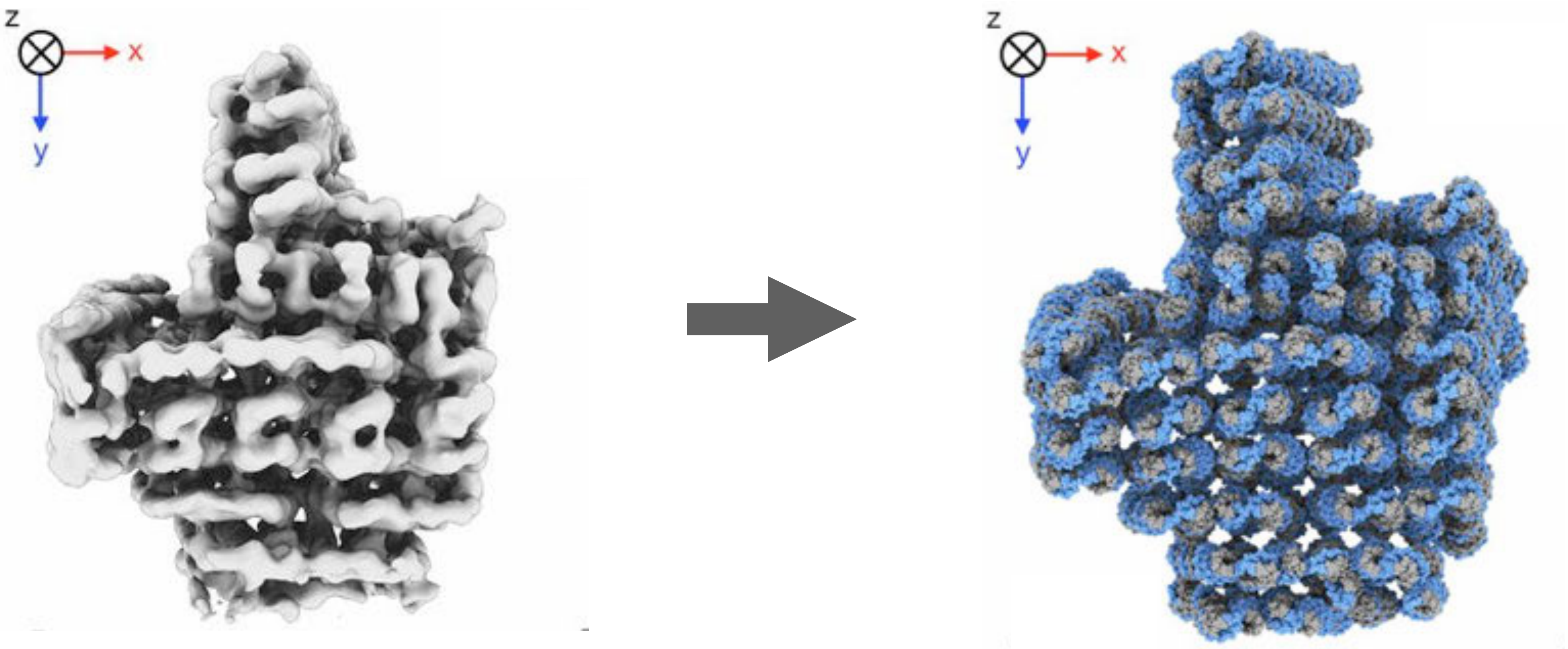
Bai *et al*, PNAS 109:20012 (2012)

Cryo-EM reconstruction versus all-atom simulation



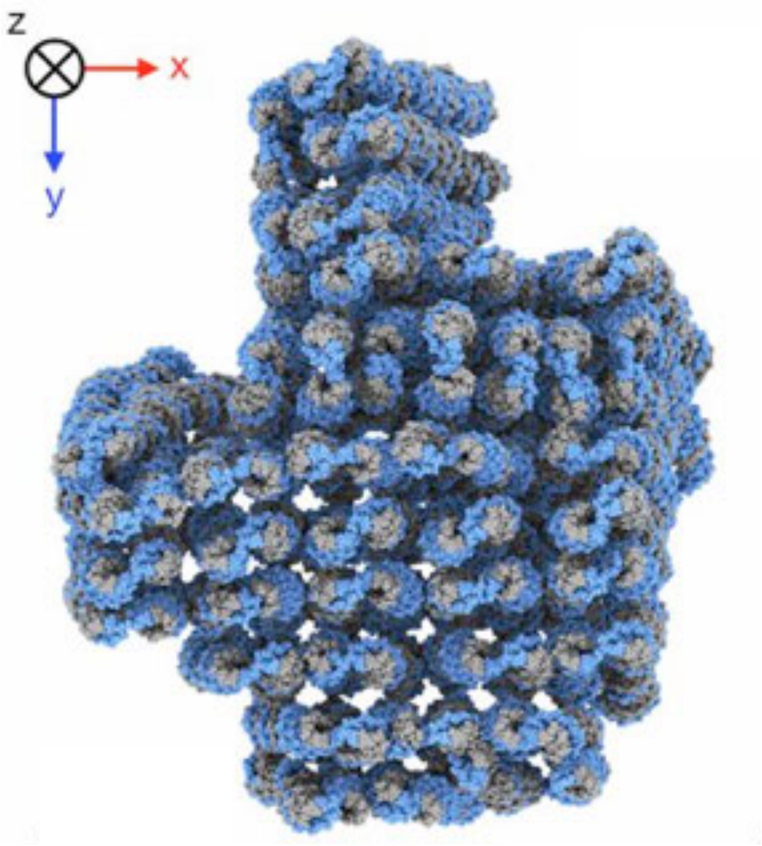
Bai *et al*, PNAS 109:20012 (2012)

Cryo-EM reconstruction versus all-atom simulation

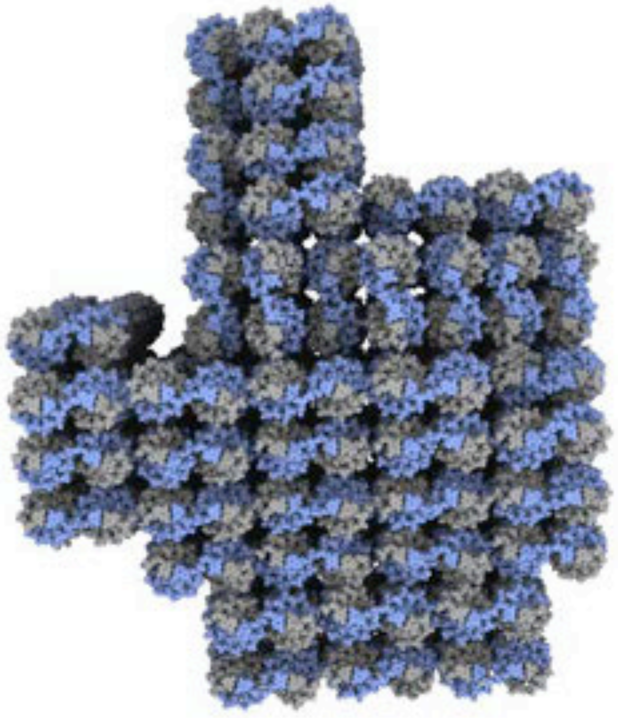


Bai *et al*, PNAS 109:20012 (2012)

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

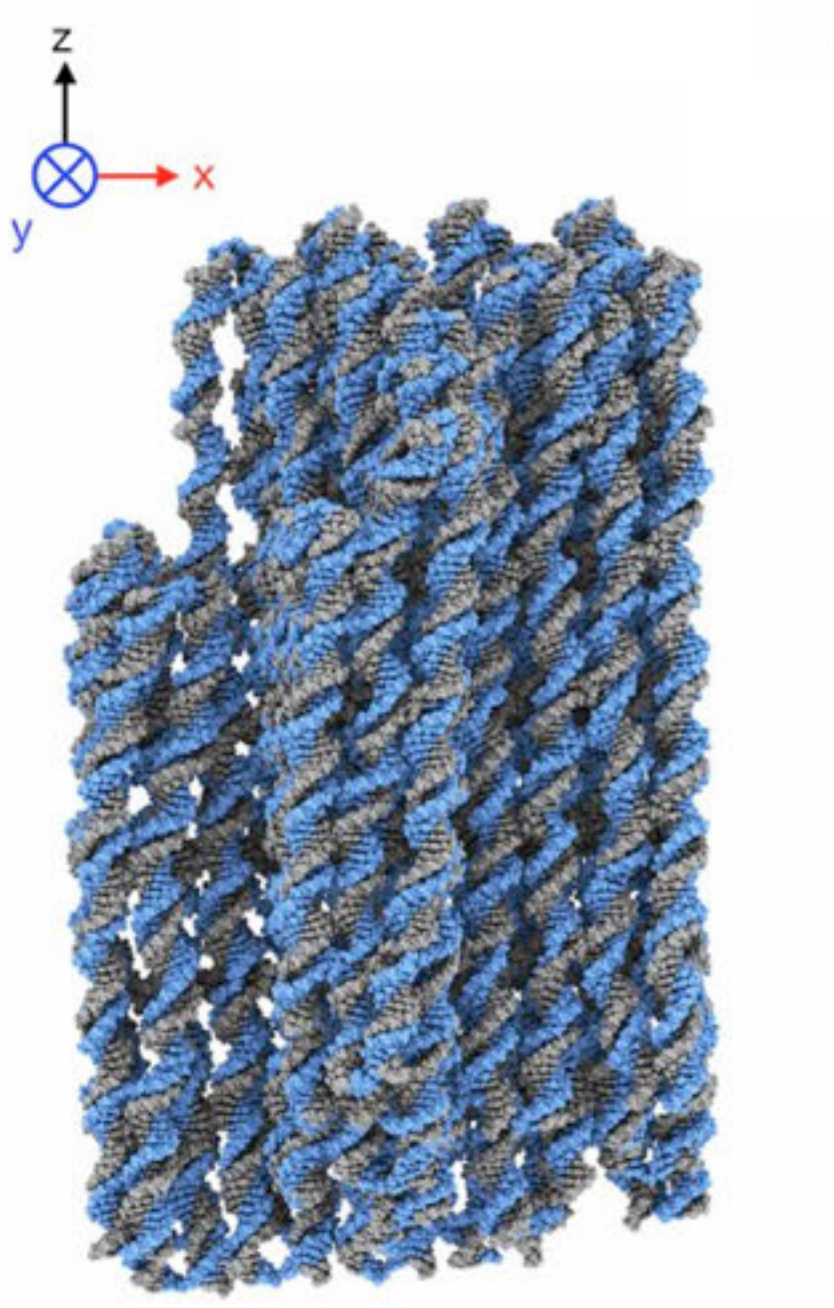


Bai *et al*, PNAS 109:20012 (2012)

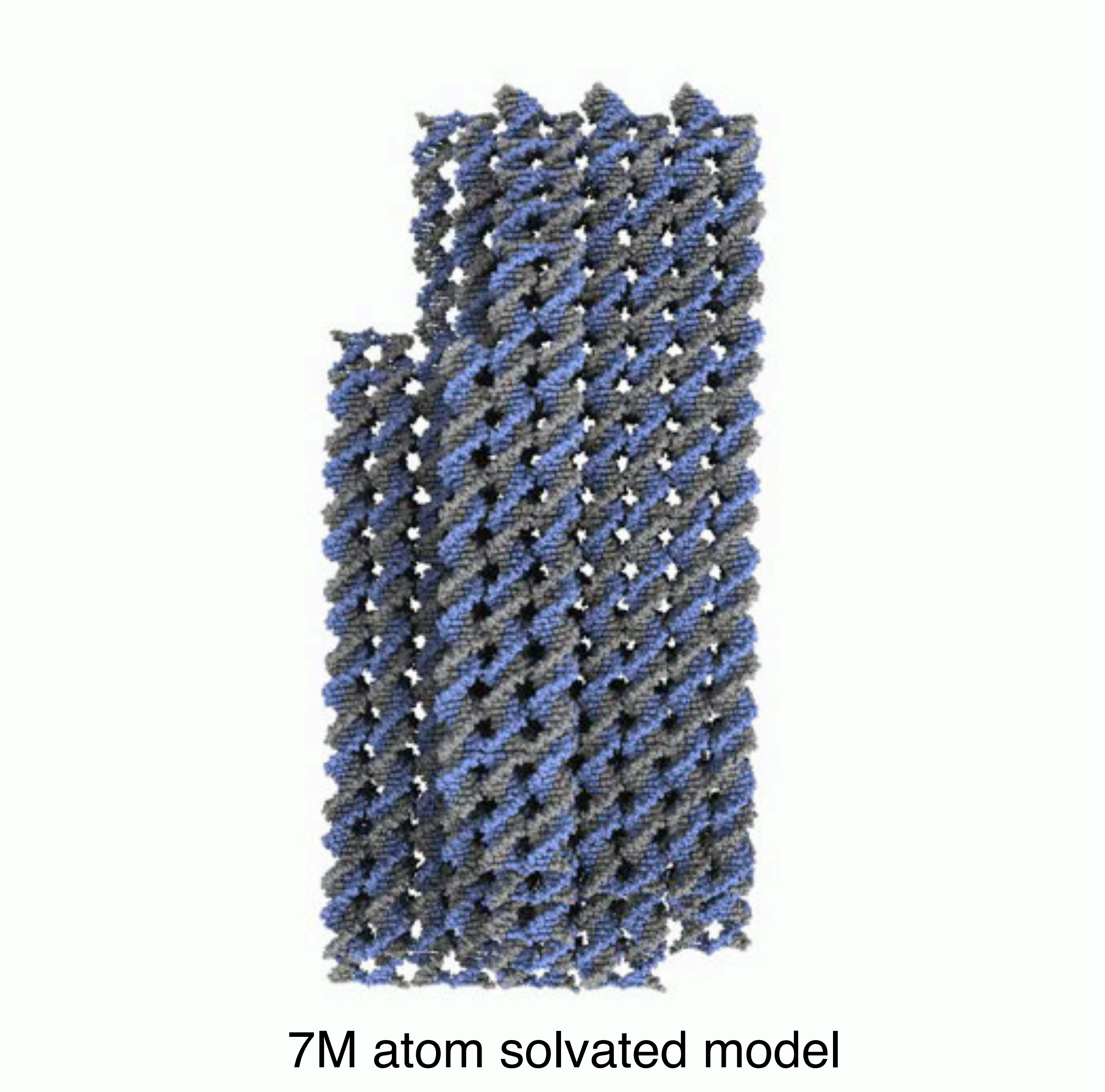


7M atom solvated model
130 ns MD trajectory

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

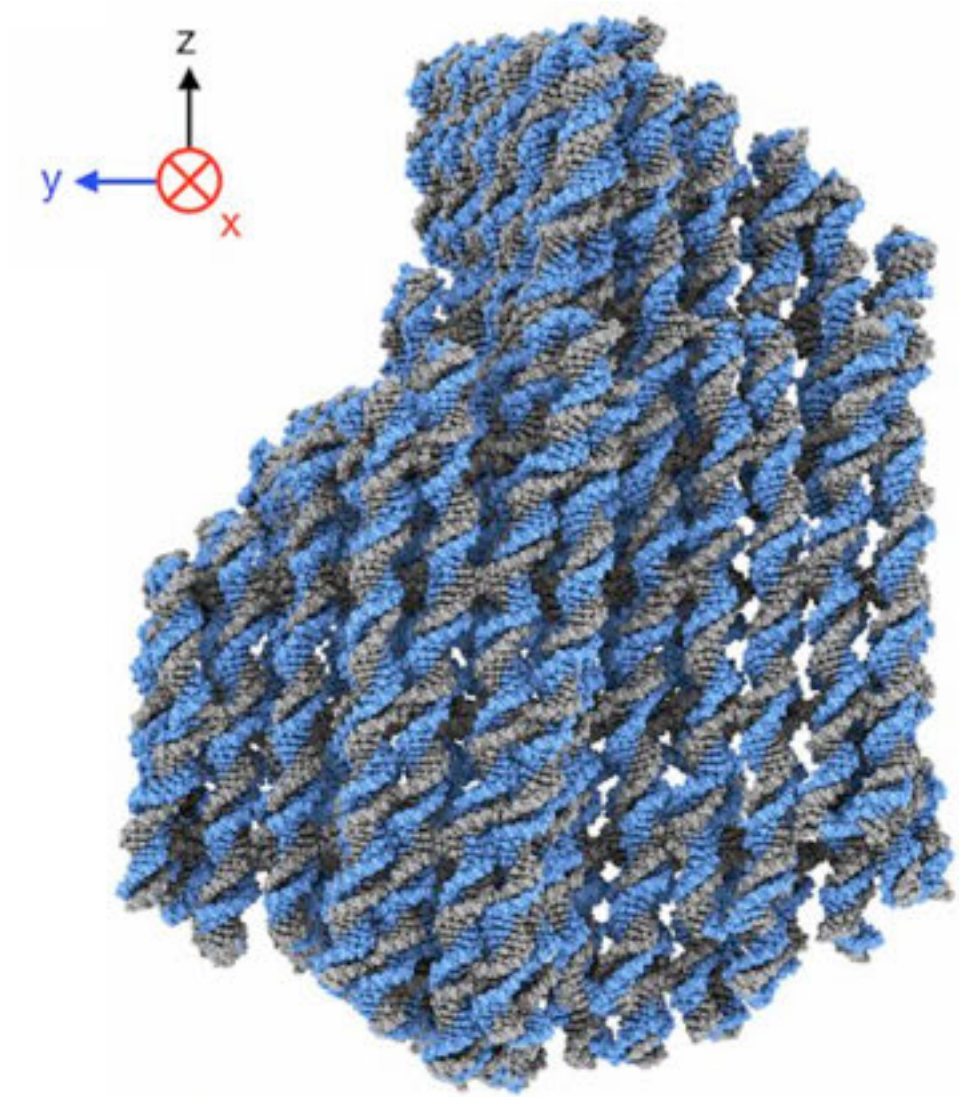


Bai *et al*, PNAS 109:20012 (2012)

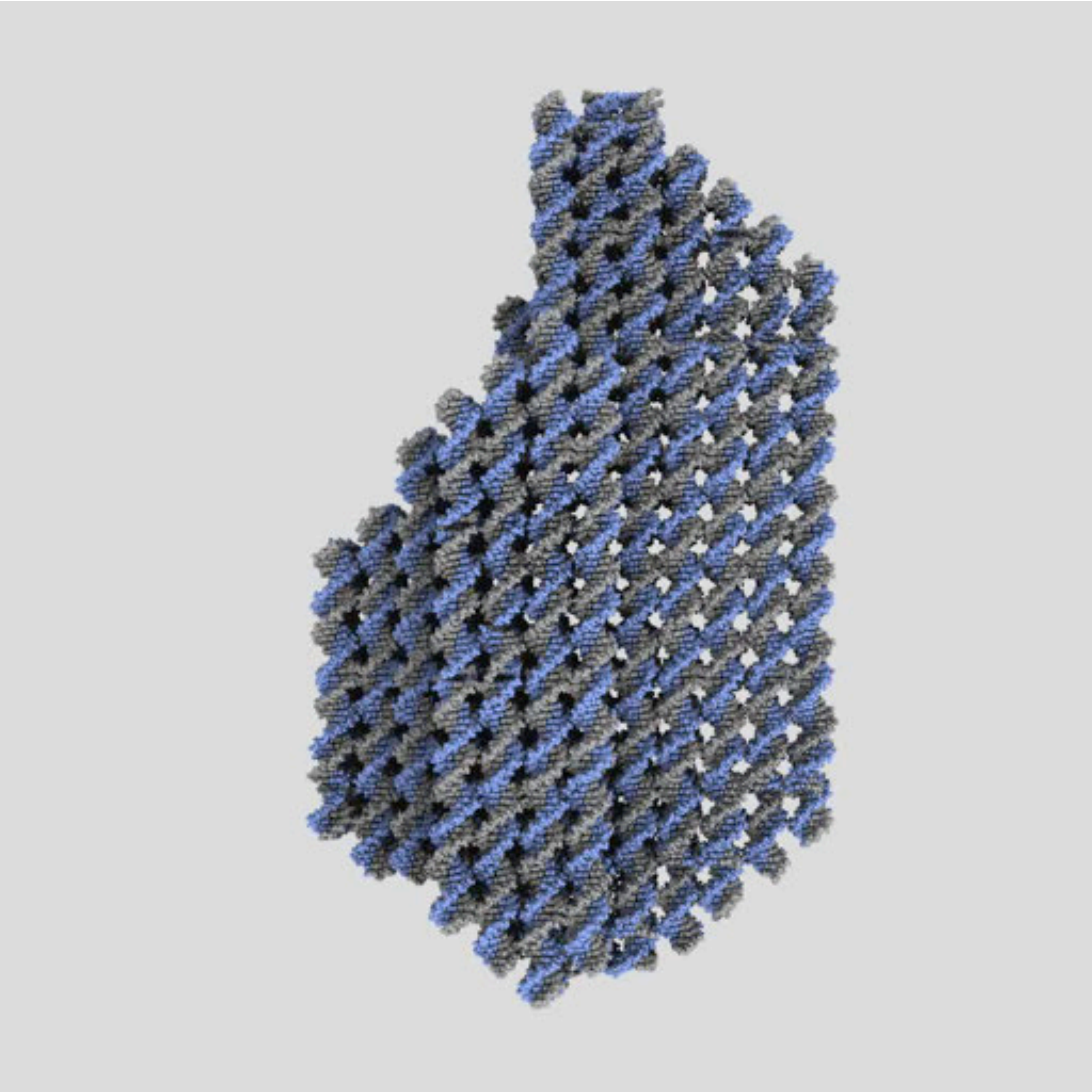


7M atom solvated model
130 ns MD trajectory

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



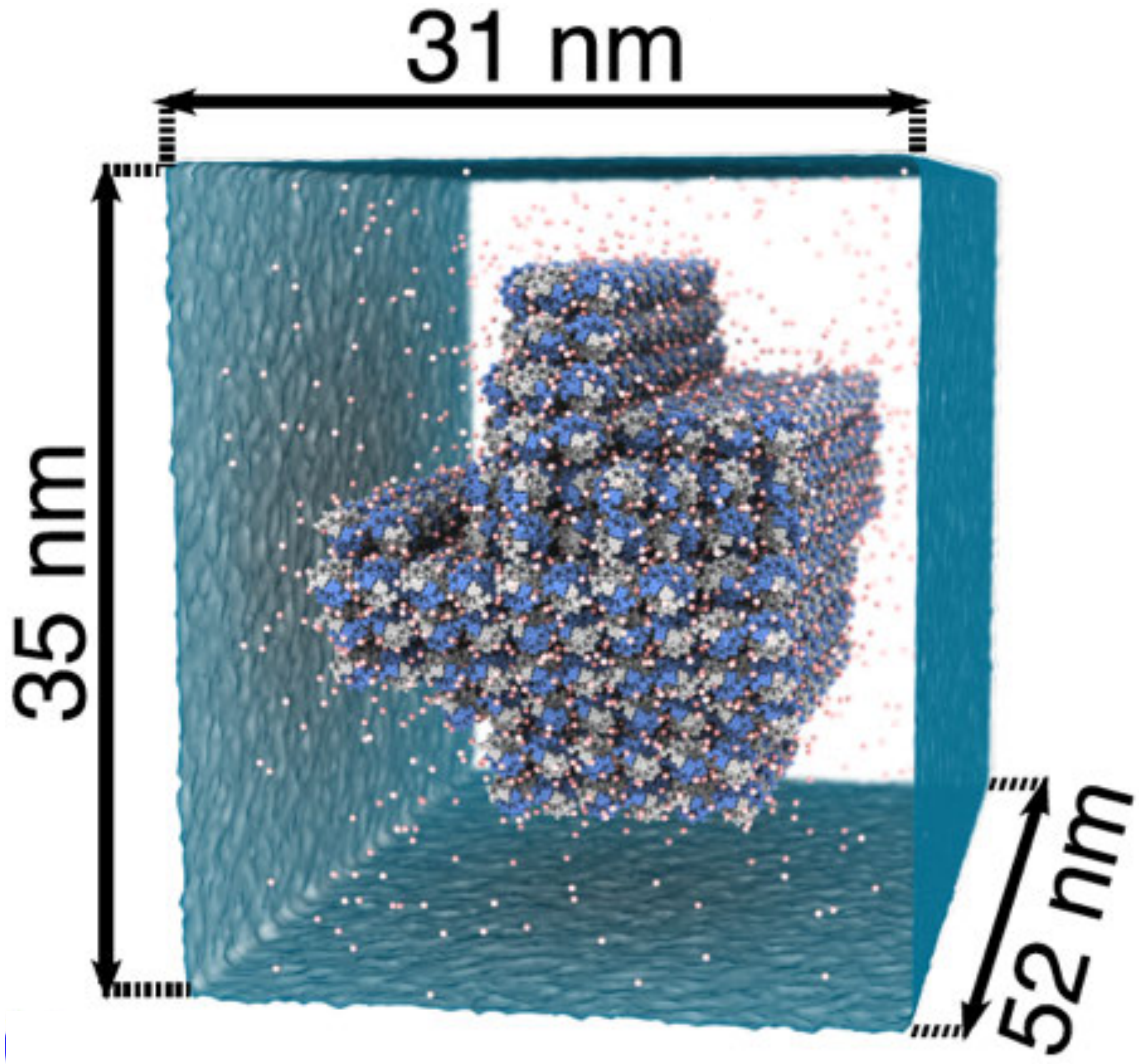
Bai *et al*, PNAS 109:20012 (2012)



7M atom solvated model
130 ns MD trajectory

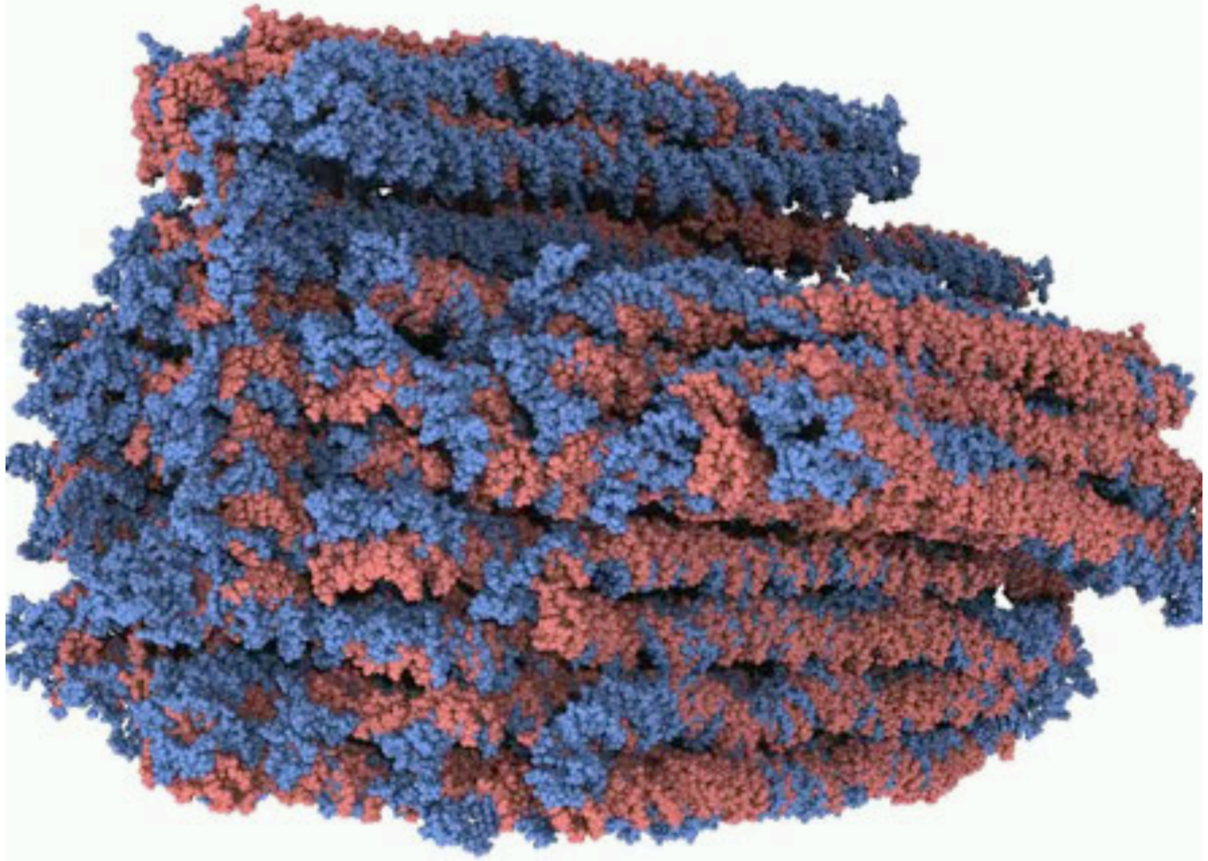
Direct comparison with cryo-EM reconstuction

Simulation on Blue Waters
(UIUC)



Time scale: 200 ns / Size ~7,000,000 atoms

Atoms move according to
classical mechanics (F= ma)

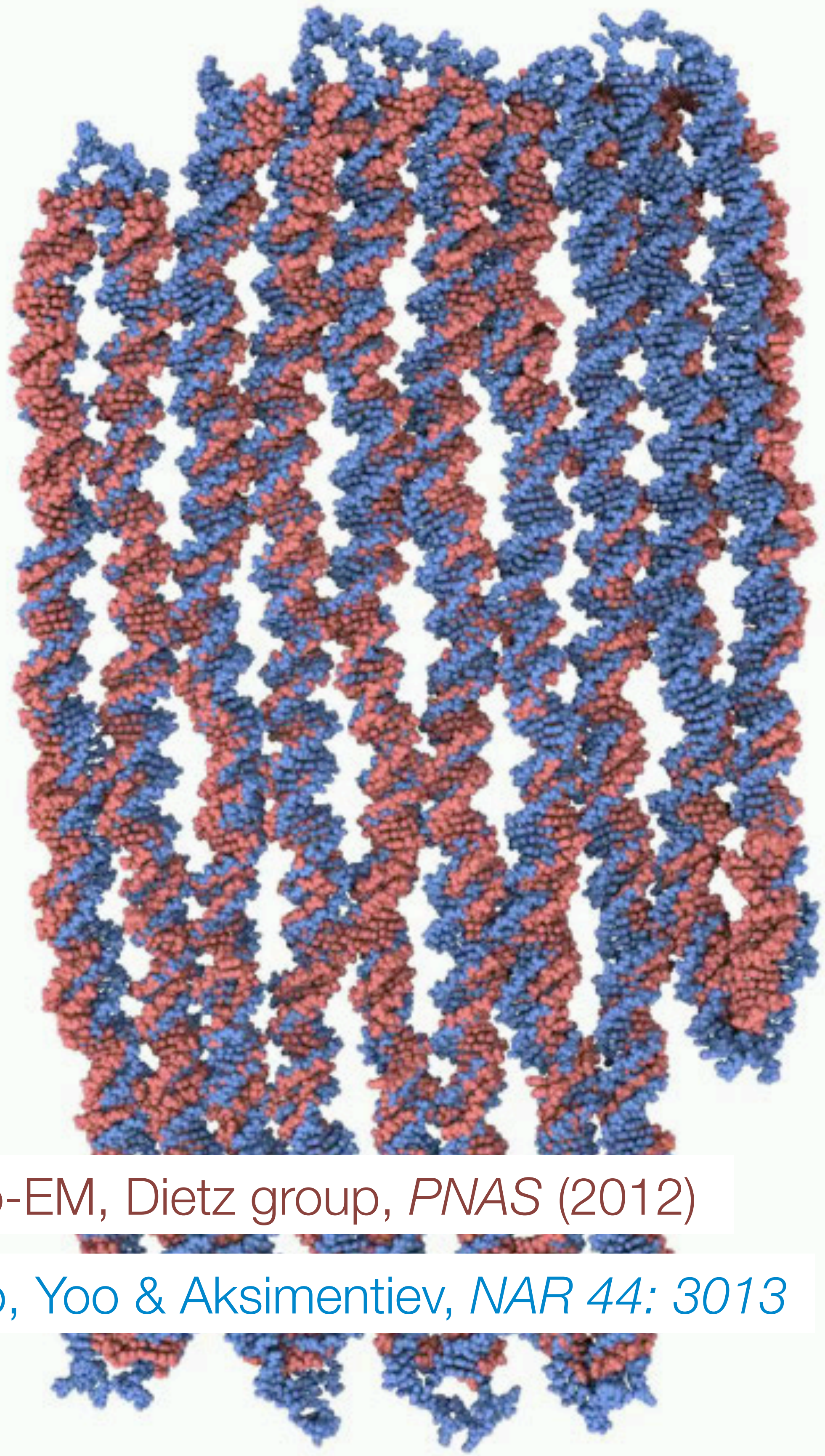


Maffeo, Yoo & Aksimentiev, *NAR 44: 3013*

Cryo-EM, Dietz group, *PNAS* (2012)

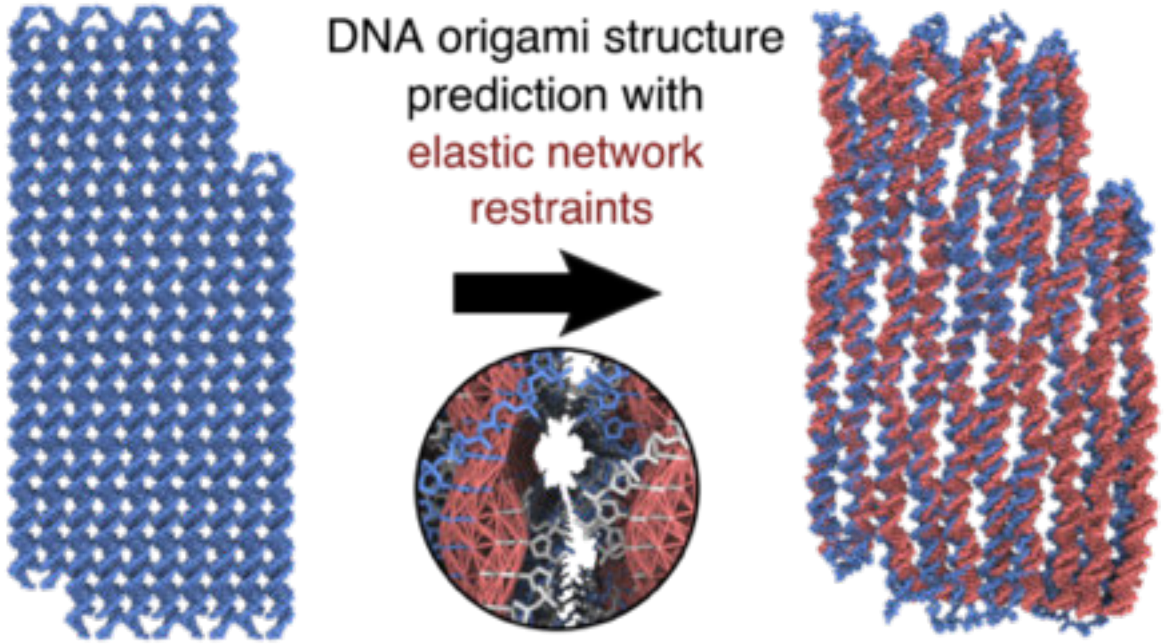
http://bionano.physics.illinois.edu/origami-structure

De novo prediction of DNA origami structure



Cryo-EM, Dietz group, *PNAS* (2012)

Maffeo, Yoo & Aksimentiev, *NAR 44: 3013*



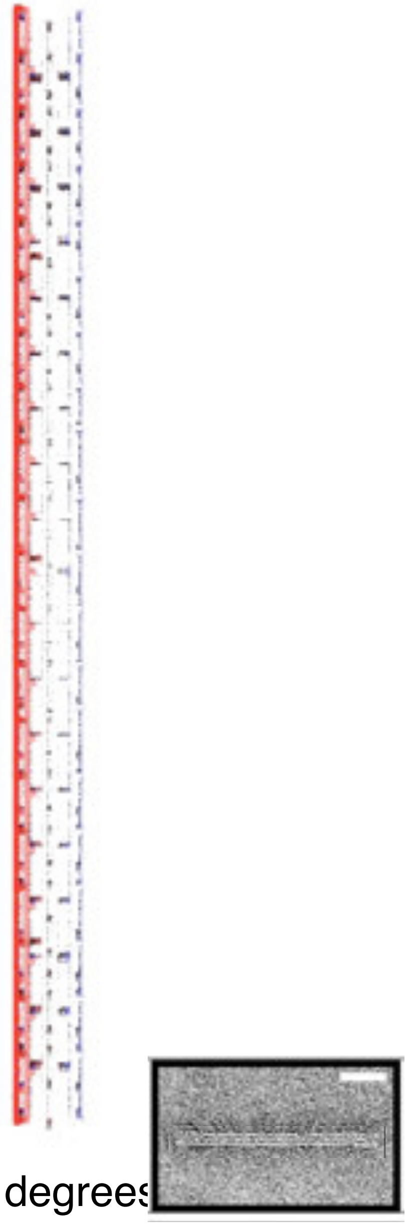
Modest computational cost
(10 hours single workstation)

Server implementation has been requested by experimentalists

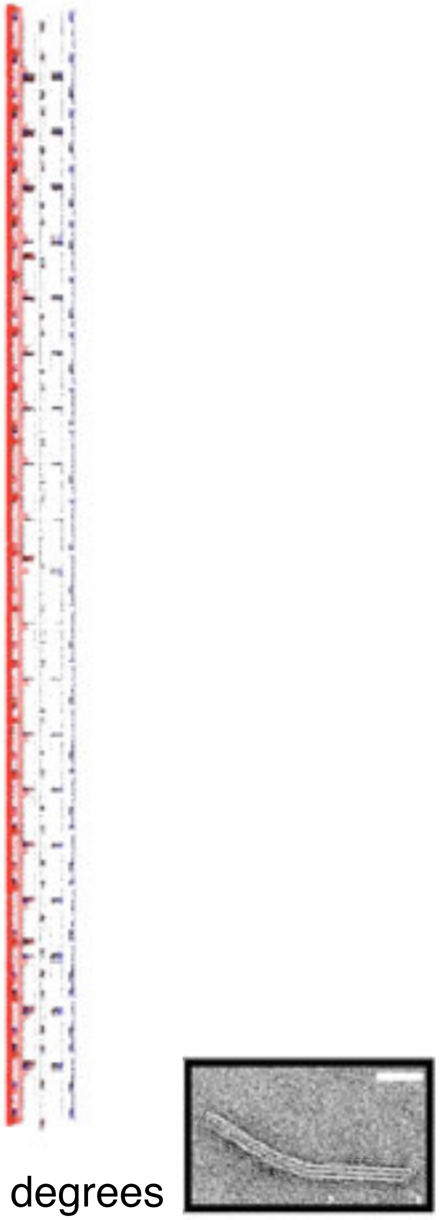
nanoHub implementation pending

http://bionano.physics.illinois.edu/origami-structure

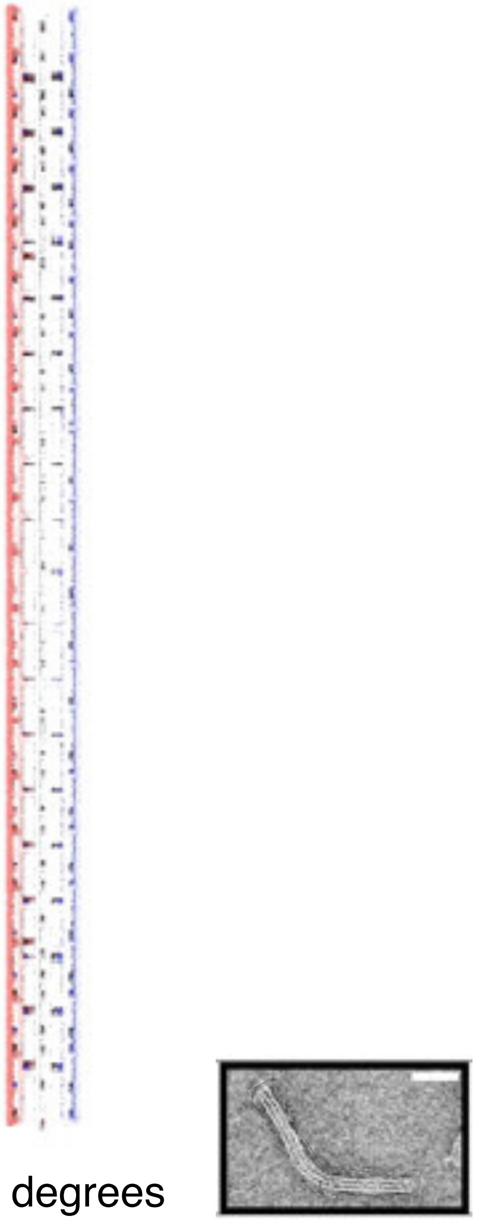
CG ENRG MD can be used routinely during the design process



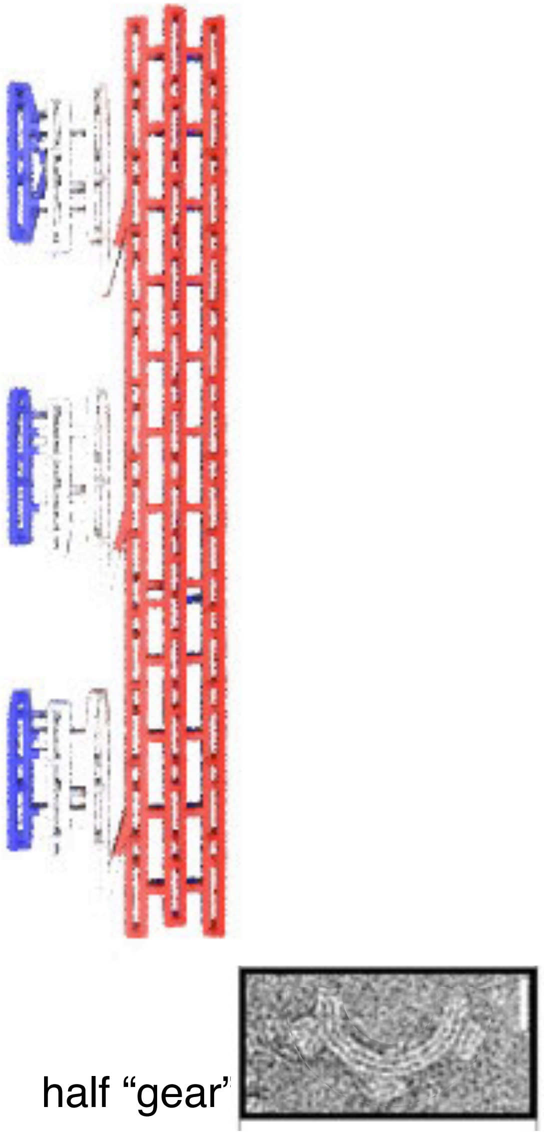
0 degrees



30 degrees



60 degrees

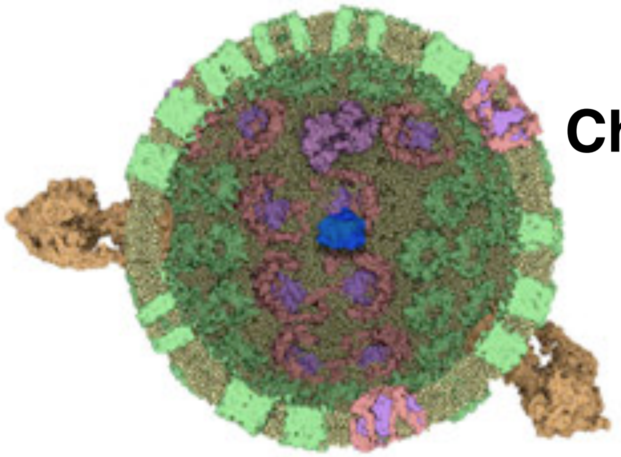


half “gear’

Modeling biomolecular systems with atomic-resolution Brownian dynamics (ARBD)

Large biological systems

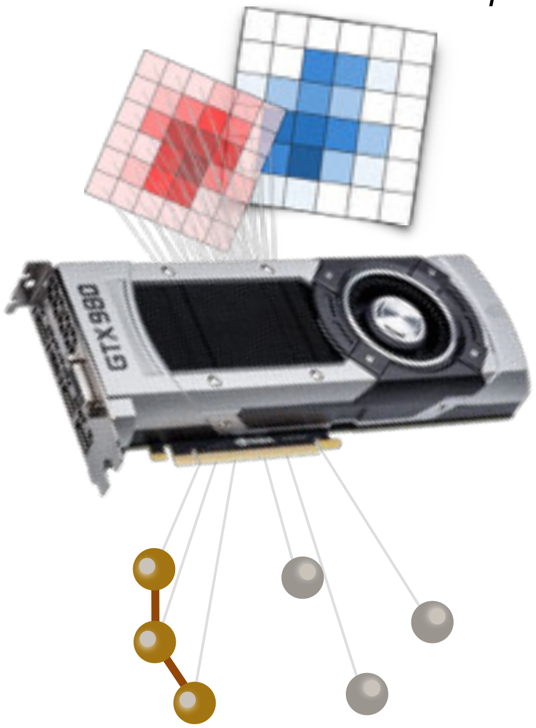
Chromatin: gene-compacting and regulating protein–DNA complex



Chromatophore: light harvesting organelle

http://www.ks.uiuc.edu/Development/Download/download.cgi?

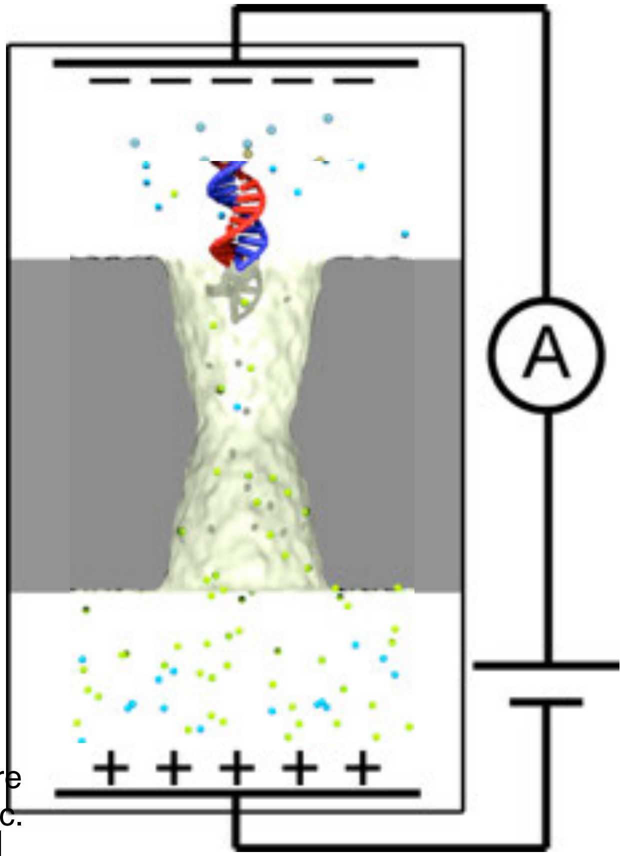
Grid-based representations of molecules with shape



Point-particle representations of polymers & small molecules

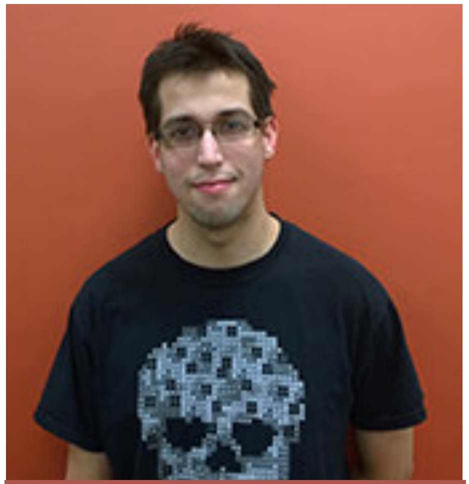
Nanotechnological systems

DNA sequencing via nanopores



Oxford Nanopore Technologies Inc.

current transients indicate passage of DNA



Michael Slone

http://bionano.physics.illinois.edu/dna-nanotechnology

Acknowledgements

NRBSC

U.S. DEPARTMENT OF ENERGY
INCITE
LEADERSHIP COMPUTING

NATIONAL INSTITUTES OF HEALTH

Jeff Comer
Maxim Belkin
Rogan Carr
David Wells
Manish Shankla
Swati Bhattacharya
Jejoong Yoo
Chen-Yu Li
Chris Maffeo
Shu-Han Chao
Derek Vandamme
Jim Wilson

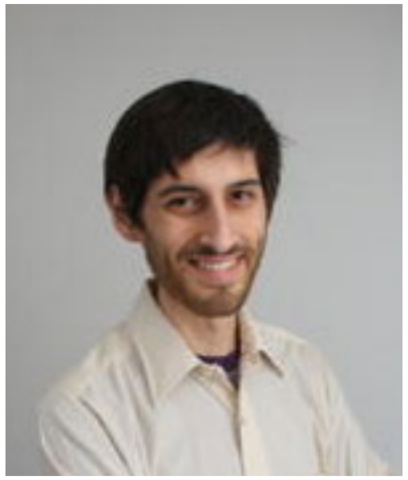

Swati Bhattacharya


Manish Shankla


Jejoong Yoo


David Wells


Chen-Yu Li


Chris Maffeo

NSF

TeraGrid™

Oxford NANOPORE Technologies™

IBM

BLUE WATERS
SUSTAINED PETASCALE COMPUTING

ACS PRF

Center for the Physics of Living Cells

US Army Corps of Engineers.
Construction Engineering Research Laboratory

ERDC DSRC
DoD Supercomputing Resource Center