On seeing and manipulating molecules

Adam Marblestone, PhD
July 2020
Talk at Oxford Future of Humanity Institute

Dynamic range of physical addressability

x, y, z ~ nanometers

X, Y, Z ~ centimeters

c ~ 4^{20} >> 23000

image vs. construct

Summary

Image:

-Using expansion microscopy and in-situ optical sequencing of DNA, we are on path to "*arbitrary resolution, infinite color microscopy*" for bio

-Sequencing neuron-indexing RNA barcodes this way should allow *full mammalian brain connectomes for ~$30M*

Construct:

-DNA origami allows ~5 nm patterning of diverse materials over ~100 nm length scales by self-assembly, with trivial/facile design

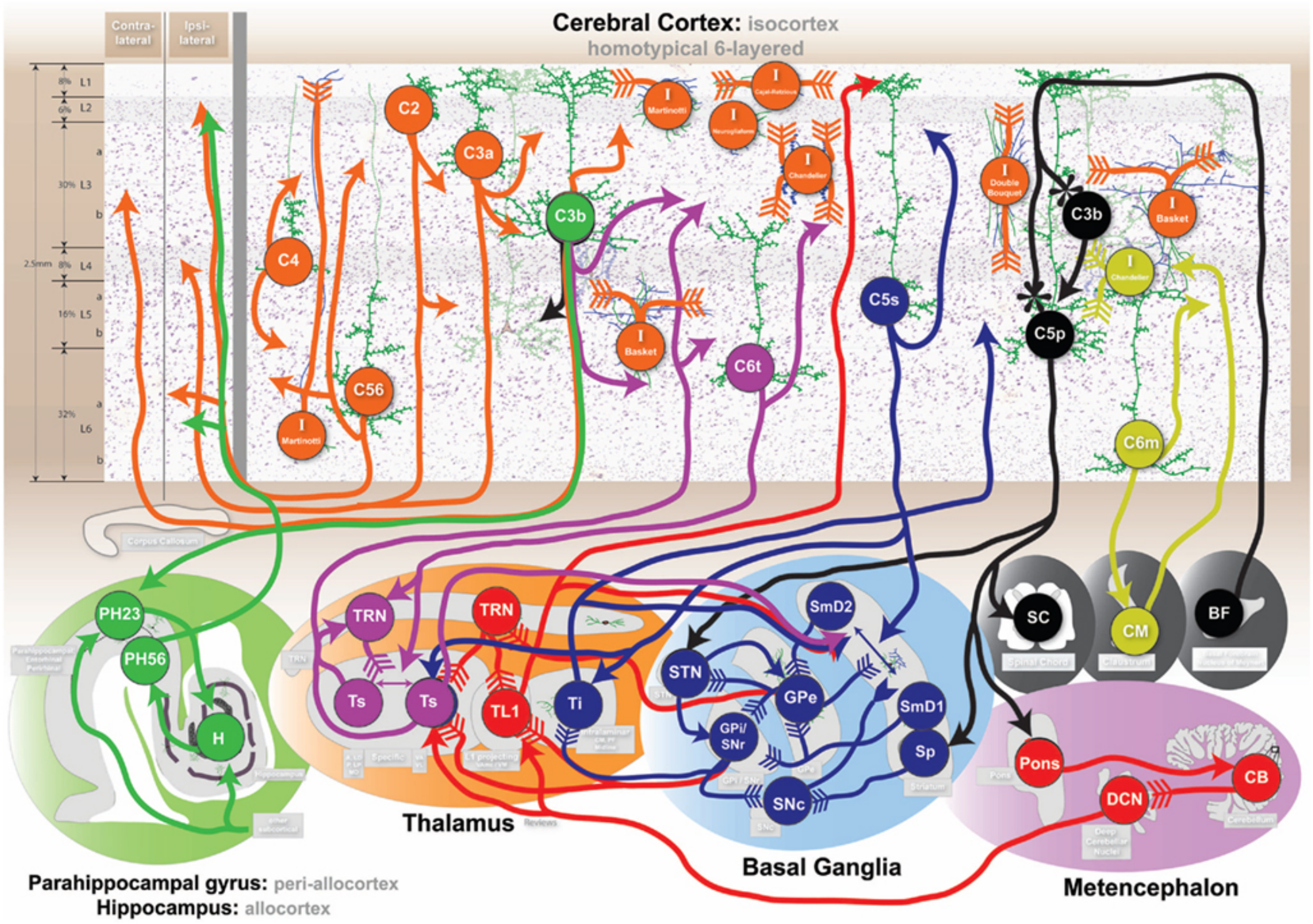
-We should be able to organize DNA origami into larger systems by patterning them on *shrunken* DNA microarrays

-A "*molecular 3D printer*", made through various steps involving DNA origami, DNA templating of proteins / peptides, and modular protein engineering with fixed backbones and *programmable interfaces*, could be an interesting first step to explore and demo "APM" principles

Part 1: image

How to scalably map the brain's architecture

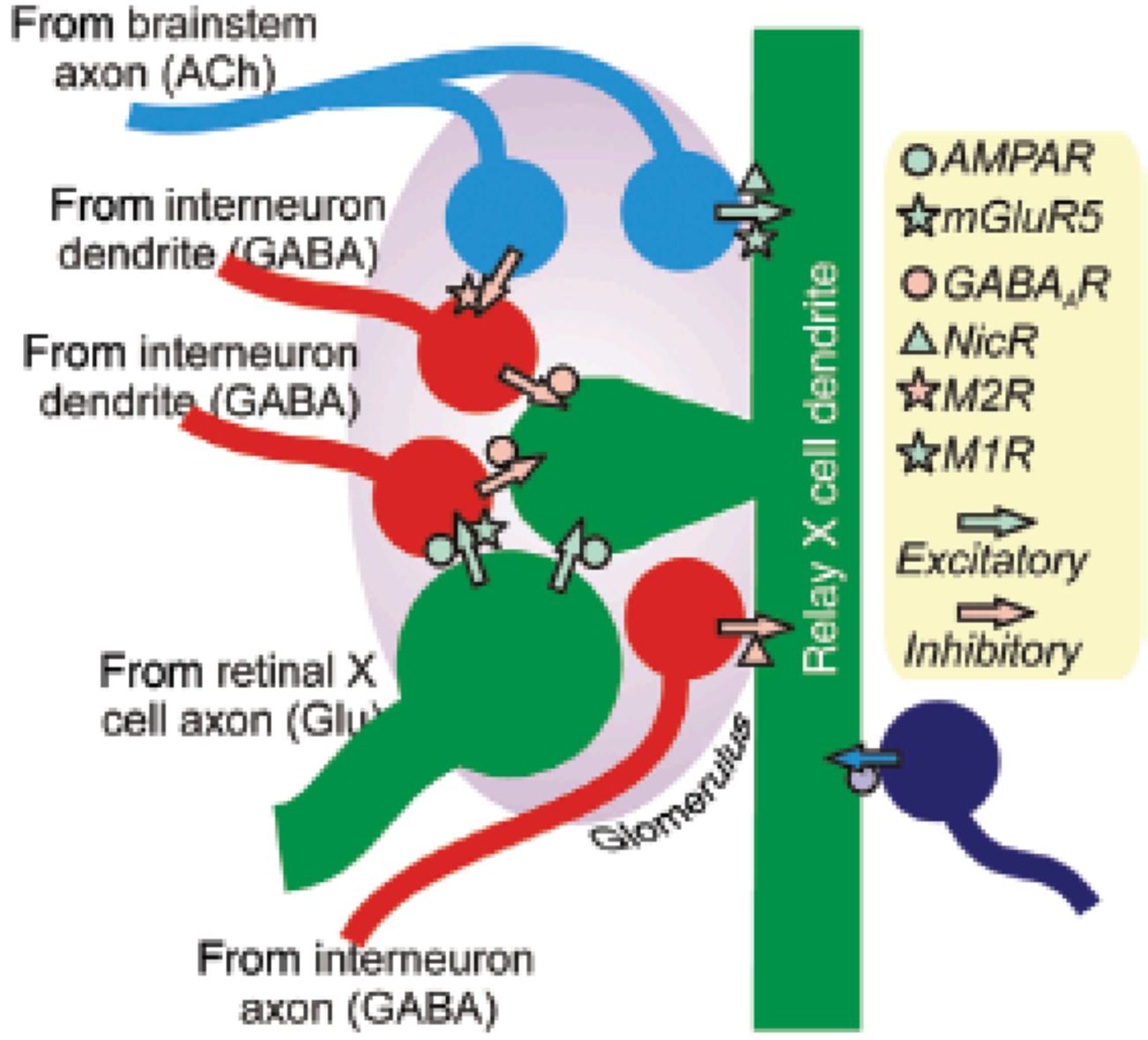
Key circuits span the entire brain (cm)



Cognitive consilience: primate non-primary neuroanatomical circuits underlying cognition

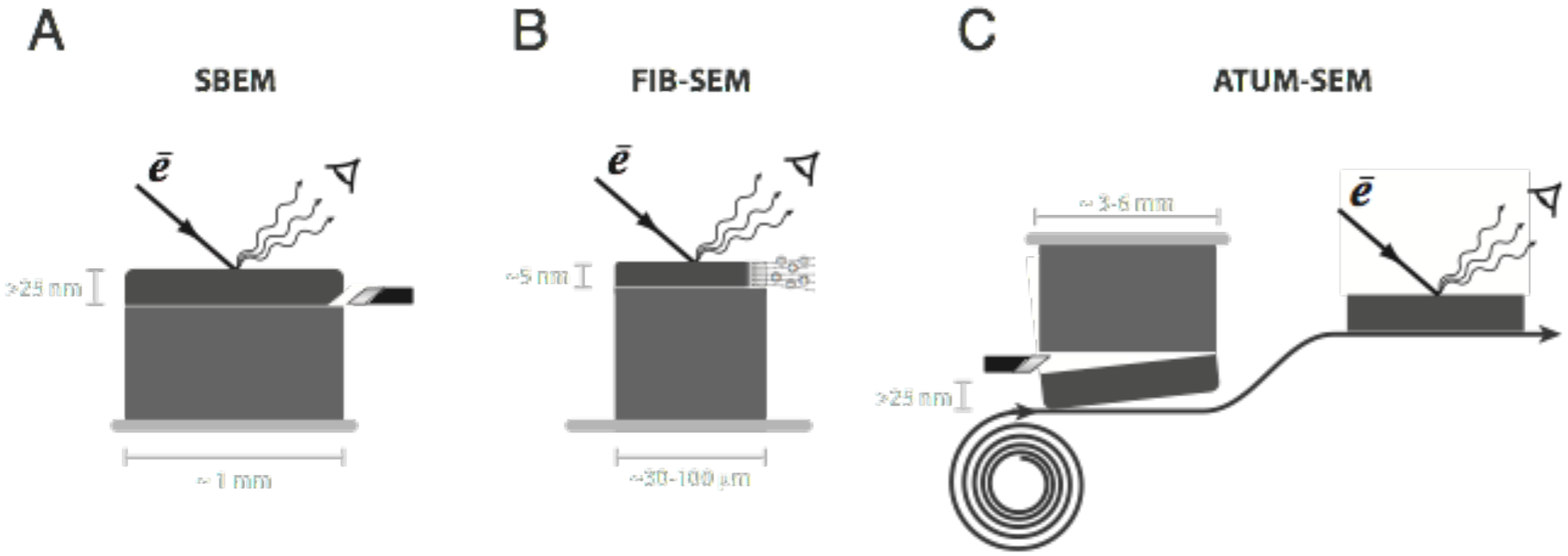
Soren Van Hout Solari[1,2]* and Rich Stoner[3]*

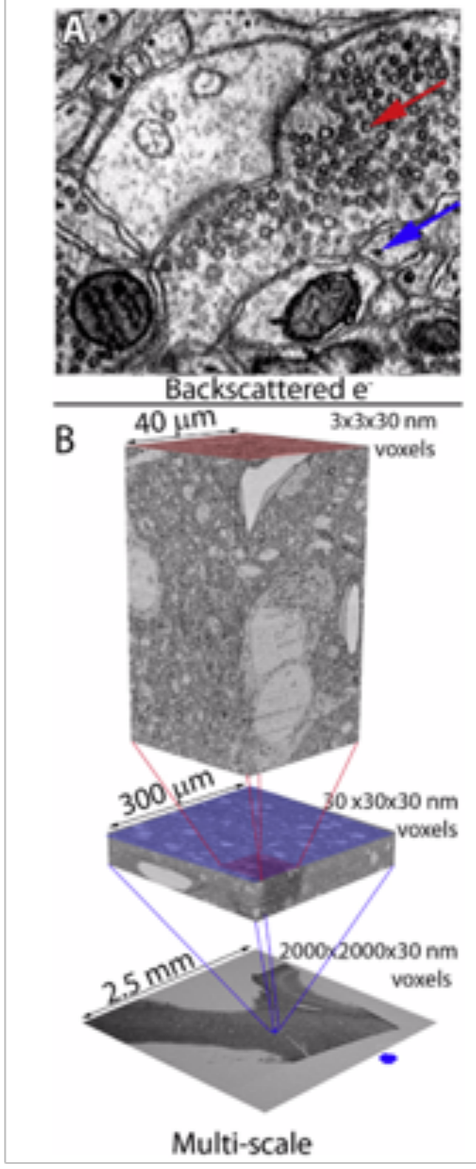
…yet are also organized at micron and nm scale



Murray Sherman

Structural brain mapping today





Current practice: widely believed that <u>electron microscopy</u> (EM) is the **only** viable method for cellular-resolution connectomics

Problems:

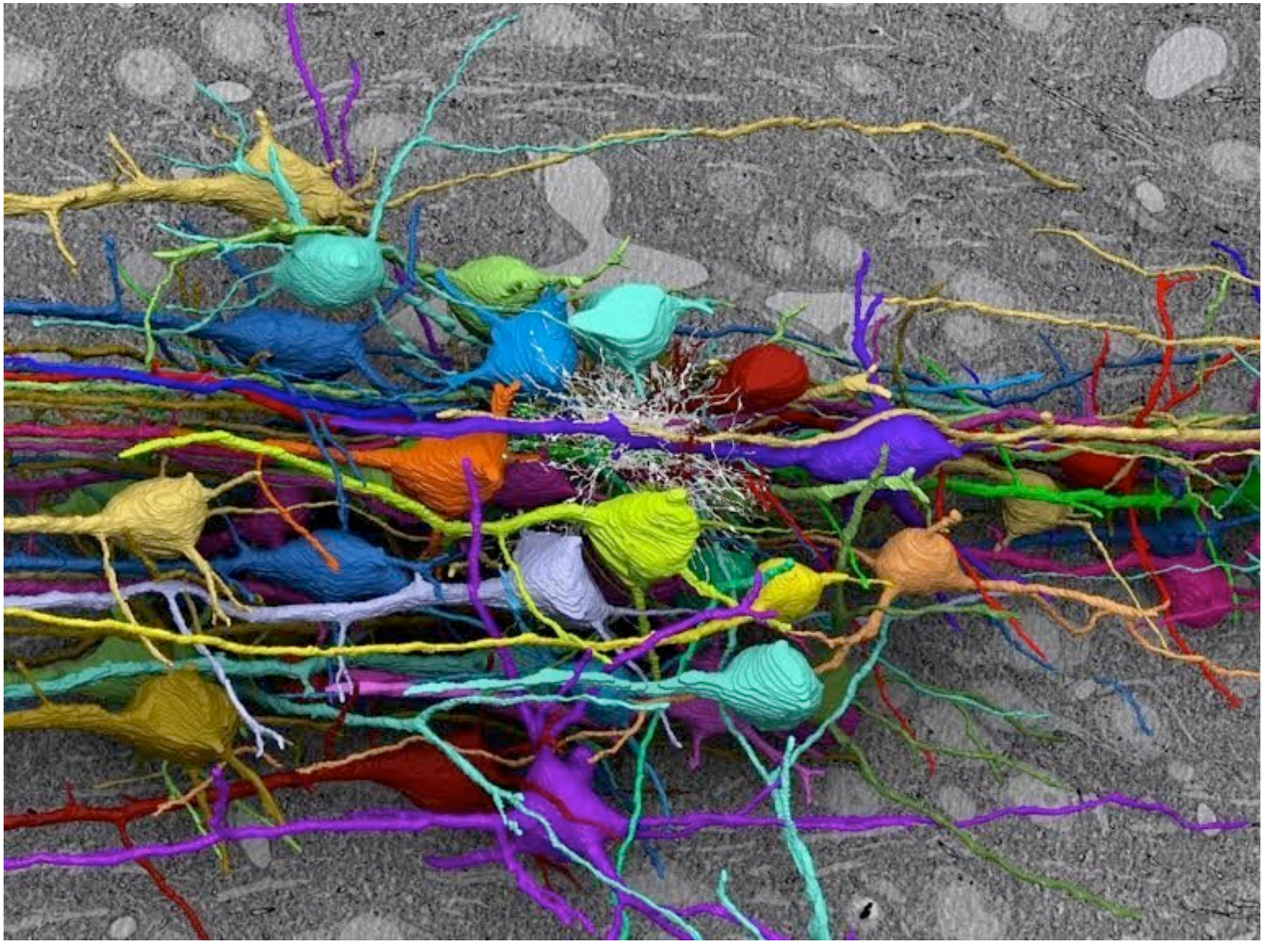
-Manual human segmentation, taking minutes per cubic *micron* (Kasthuri, 2015)

-Does not provide molecular information

x,y,z < 10 nm

X,Y,Z < 1 mm

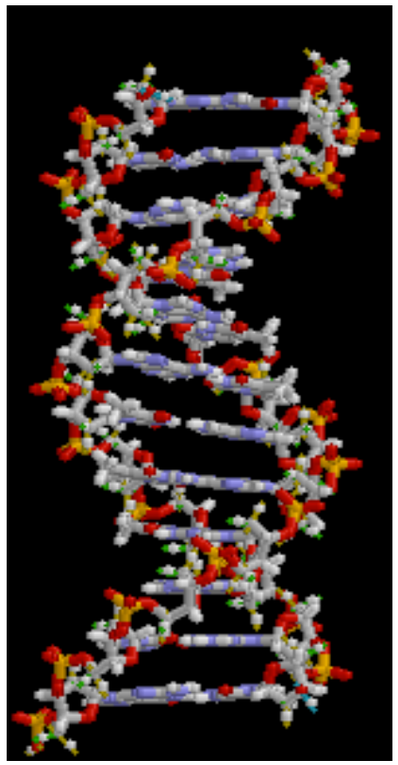
c ~ 1



Jeff Lichtman

Scalable brain mapping: unique ID's for structure

4^N possible DNA sequences of length N "letters"



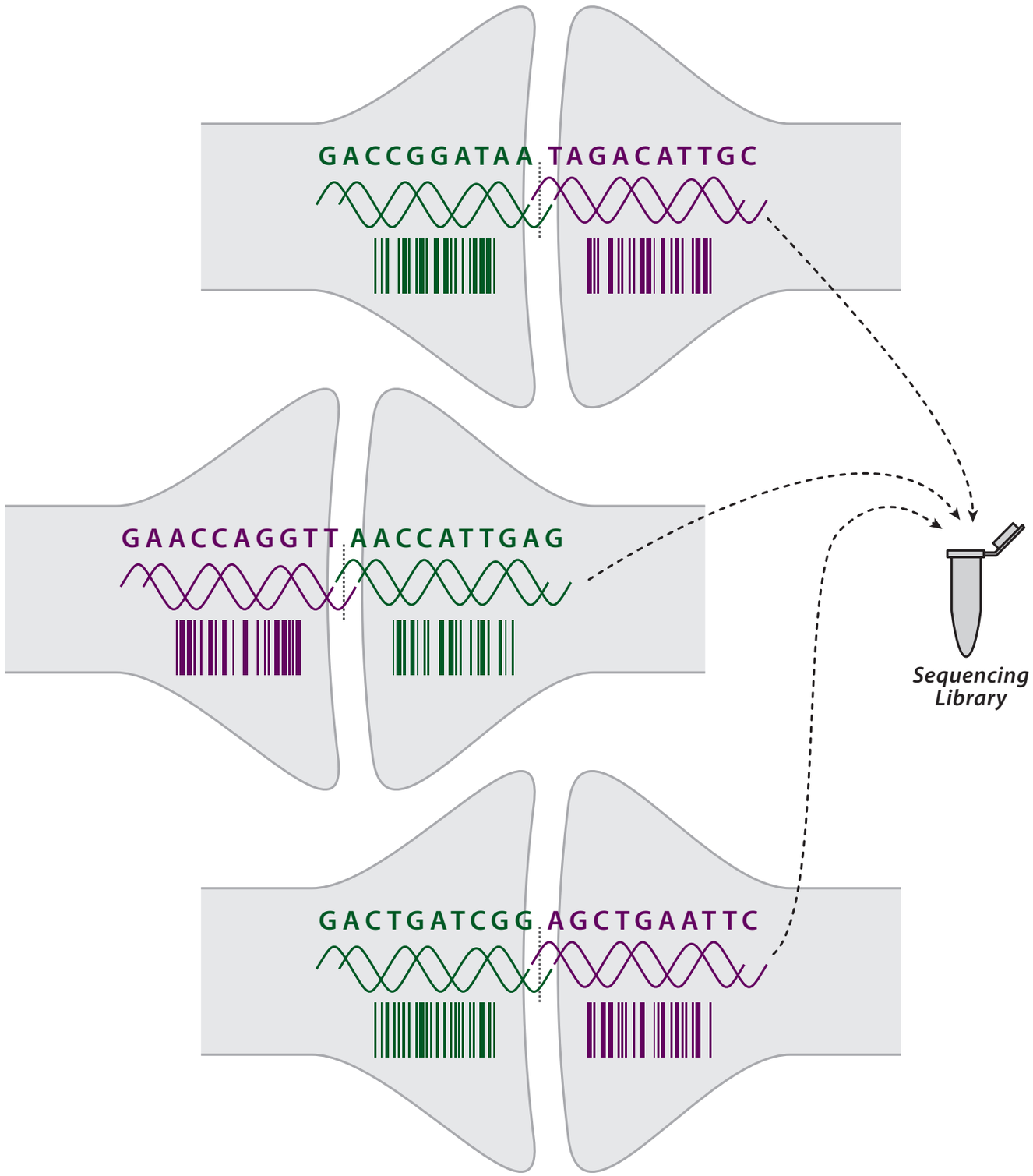
= 0I0I00I0I0I0I0I00I0I0I
0I0I0I0I0I00I0I0II0III
00I0II0I0I0I0I1I0I00II

=

Zador, Cepko, Tabin, Walsh, Church et al:

can give every neuron a uniquely-identifiable DNA "barcode"

DNA Barcodes via Bulk Sequencing



Zador lab

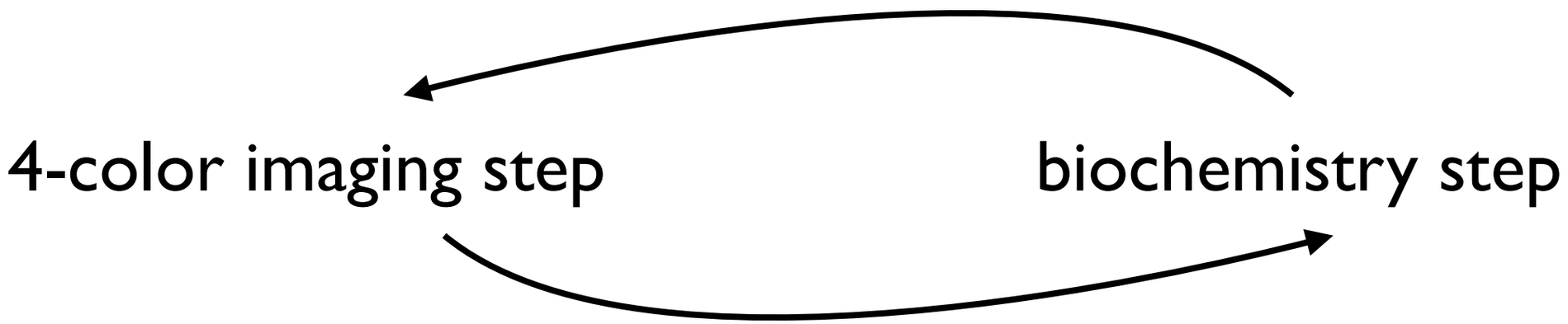
BOINC

Connectivity Matrix

Scalable brain mapping: unique ID's for structure

digital **4^Ncolor** microscopy

# of cycles	*# of molecules discriminated*
1	4
10	4^{10}, about 10^6
30	4^{30}, about 10^{18}



... N cycles

Read one letter of RNA at a time, *sequentially*

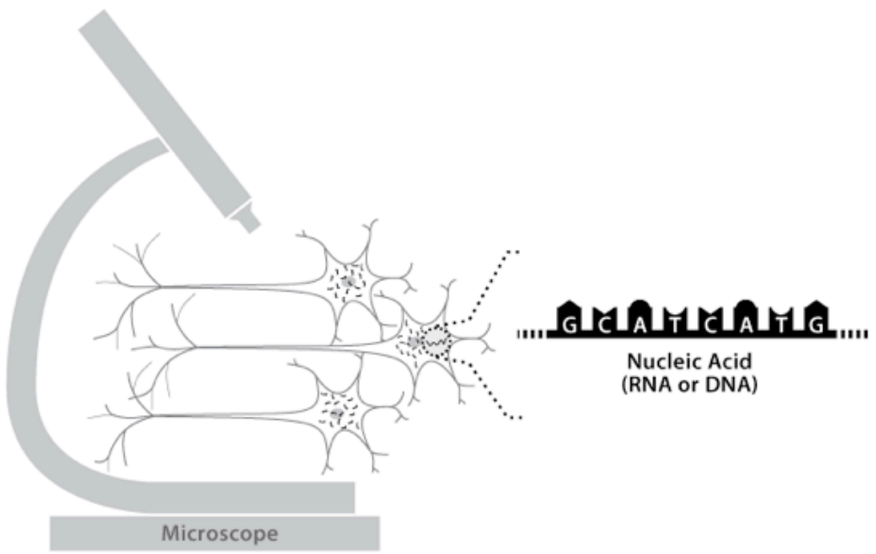
Church lab circa ~2014

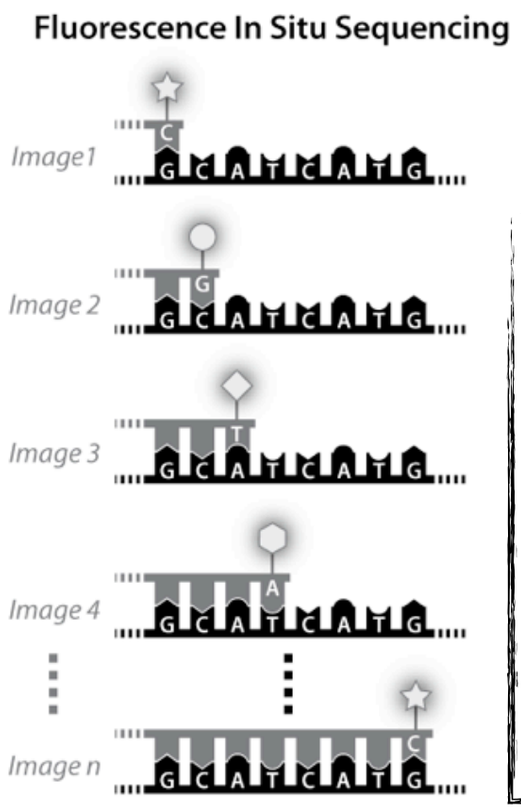
digital 4^N**color** microscopy

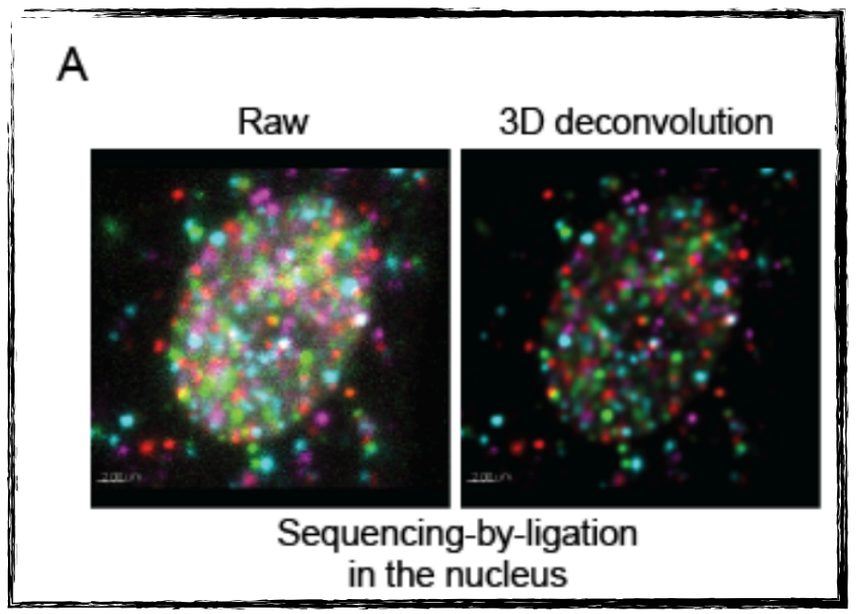
# of cycles	*# of molecules discriminated*
1	4
10	4^{10}, about 10^6
30	4^{30}, about 10^{18}

Fluorescent In-Situ RNA Sequencing (FISSEQ):

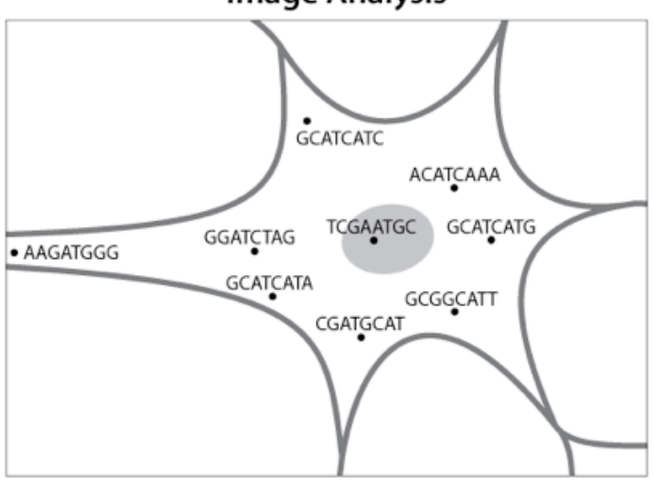
A







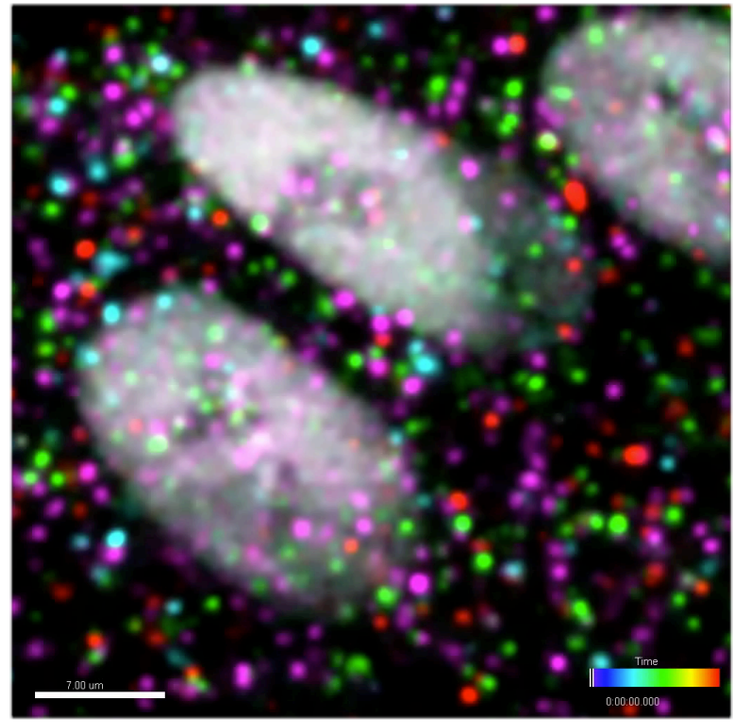
B



Highly Multiplexed Subcellular RNA Sequencing in Situ

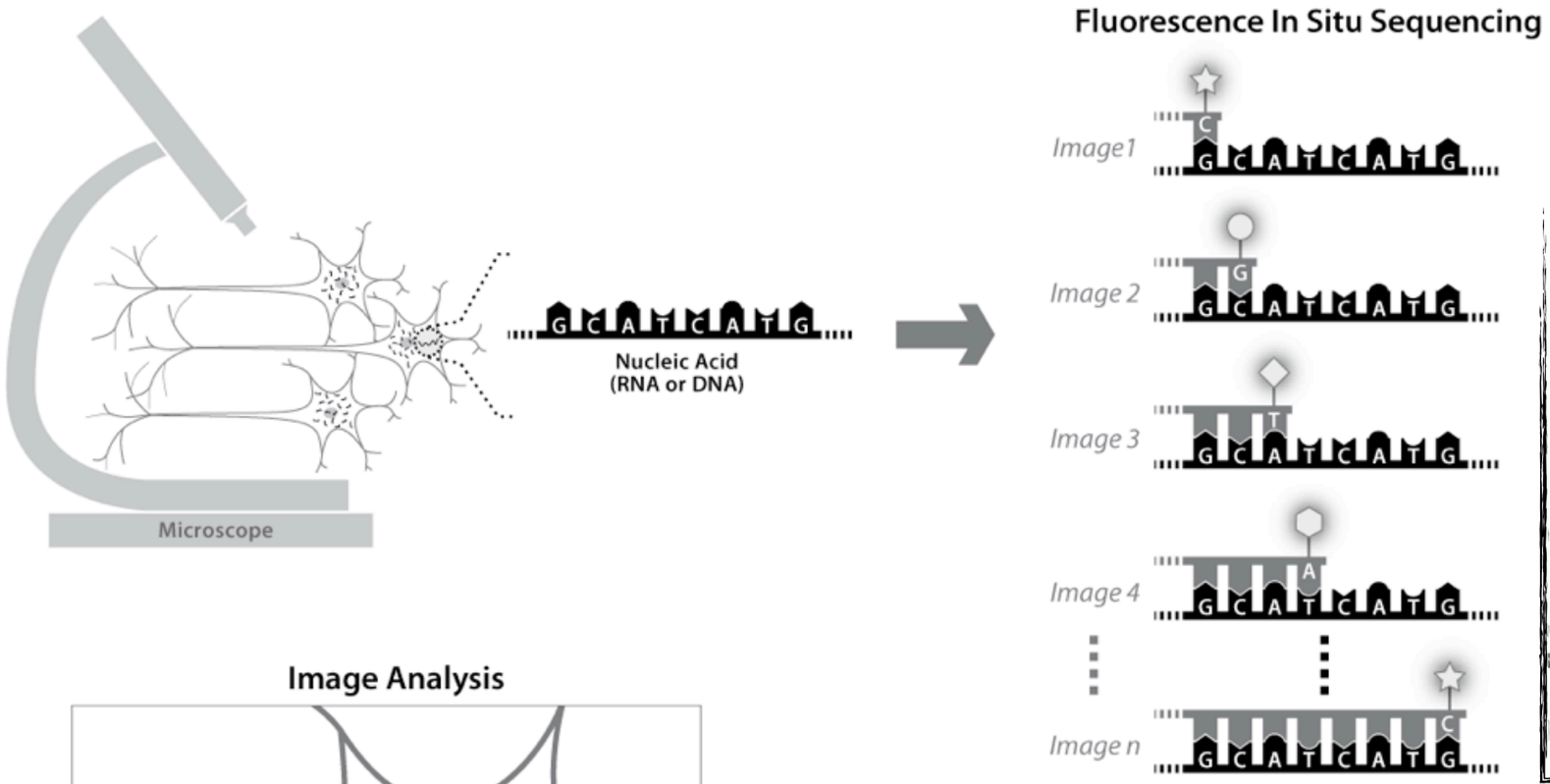
Je Hyuk Lee,[1,2]*† Evan R. Daugharthy,[1,2,4]* Jonathan Scheiman,[1,2] Reza Kalhor,[2] Joyce L. Yang,[2] Thomas C. Ferrante,[1] Richard Terry,[1] Sauveur S. F. Jeanty,[1] Chao Li,[1] Ryoji Amamoto,[3] Derek T. Peters,[3] Brian M. Turczyk,[1] Adam H. Marblestone,[1,2] Samuel A. Inverso,[1] Amy Bernard,[5] Prashant Mali,[2] Xavier Rios,[2] John Aach,[2] George M. Church[1,2]†

digital **4^N color** microscopy

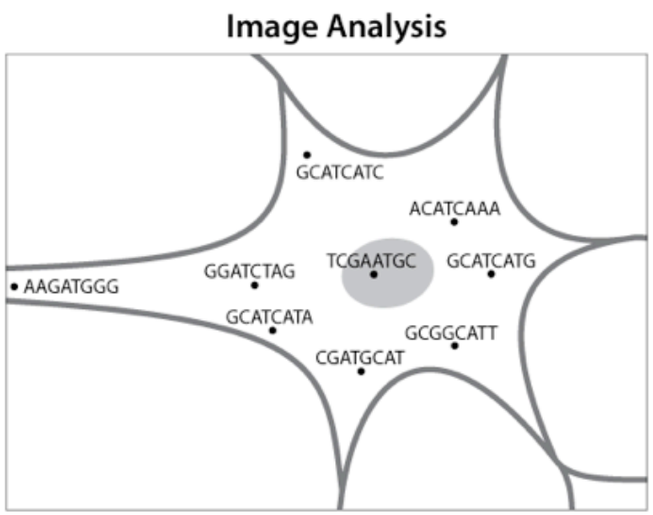


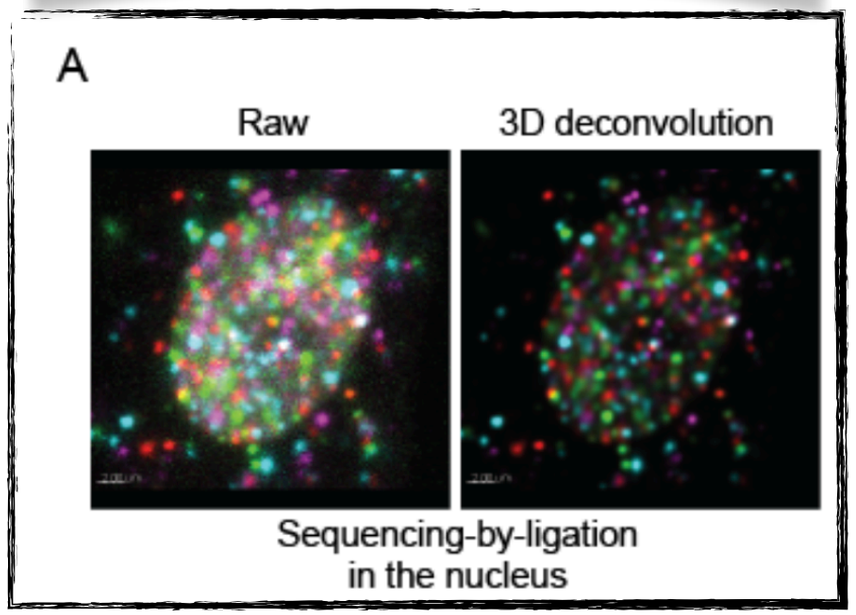
Fluorescent In-Situ RNA Sequencing (FISSEQ):

A



B

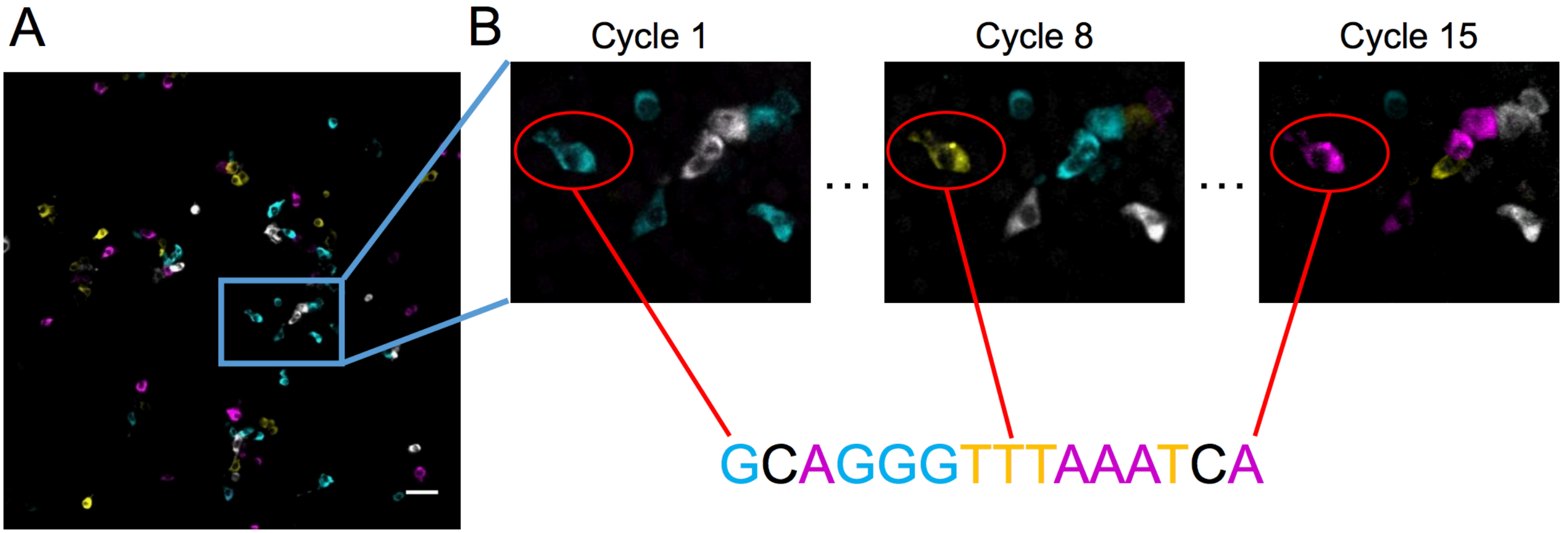




Highly Multiplexed Subcellular RNA Sequencing in Situ

Je Hyuk Lee,[1,2]*† Evan R. Daugharthy,[1,2,4]* Jonathan Scheiman,[1,2] Reza Kalhor,[2] Joyce L. Yang,[2] Thomas C. Ferrante,[1] Richard Terry,[1] Sauveur S. F. Jeanty,[1] Chao Li,[1] Ryoji Amamoto,[3] Derek T. Peters,[3] Brian M. Turczyk,[1] Adam H. Marblestone,[1,2] Samuel A. Inverso,[1] Amy Bernard,[5] Prashant Mali,[2] Xavier Rios,[2] John Aach,[2] George M. Church[1,2]†

In-situ readout of Zador barcodes



Efficient *in situ* barcode sequencing using padlock probe-based BaristaSeq

Xiaoyin Chen[1], Yu-Chi Sun[1], George M Church[2, 3], Je Hyuk Lee[1], and Anthony M Zador[1]

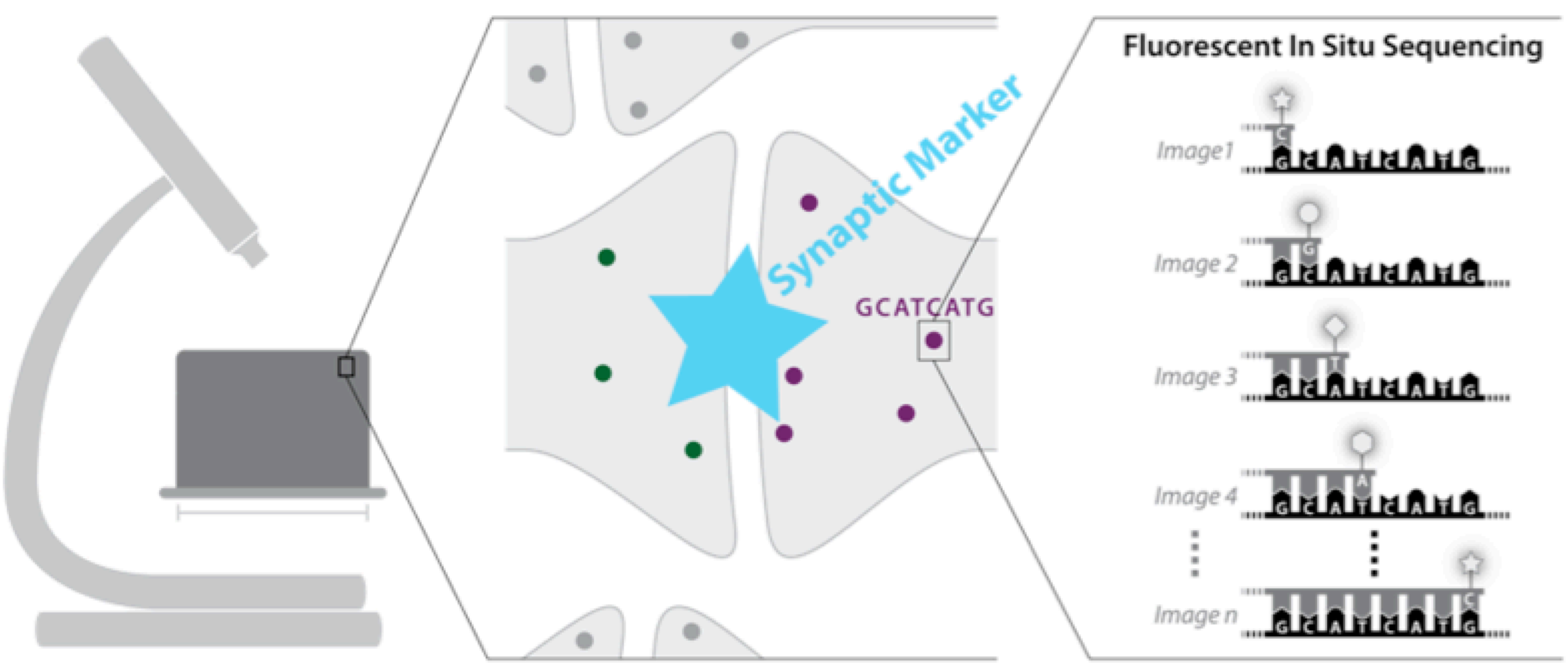
In-situ barcode connectomics

x,y,z ~ 600 nm

X,Y,Z > 1 cm

c ~ infinity

Synaptic resolution requires x,y,z < 100 nm



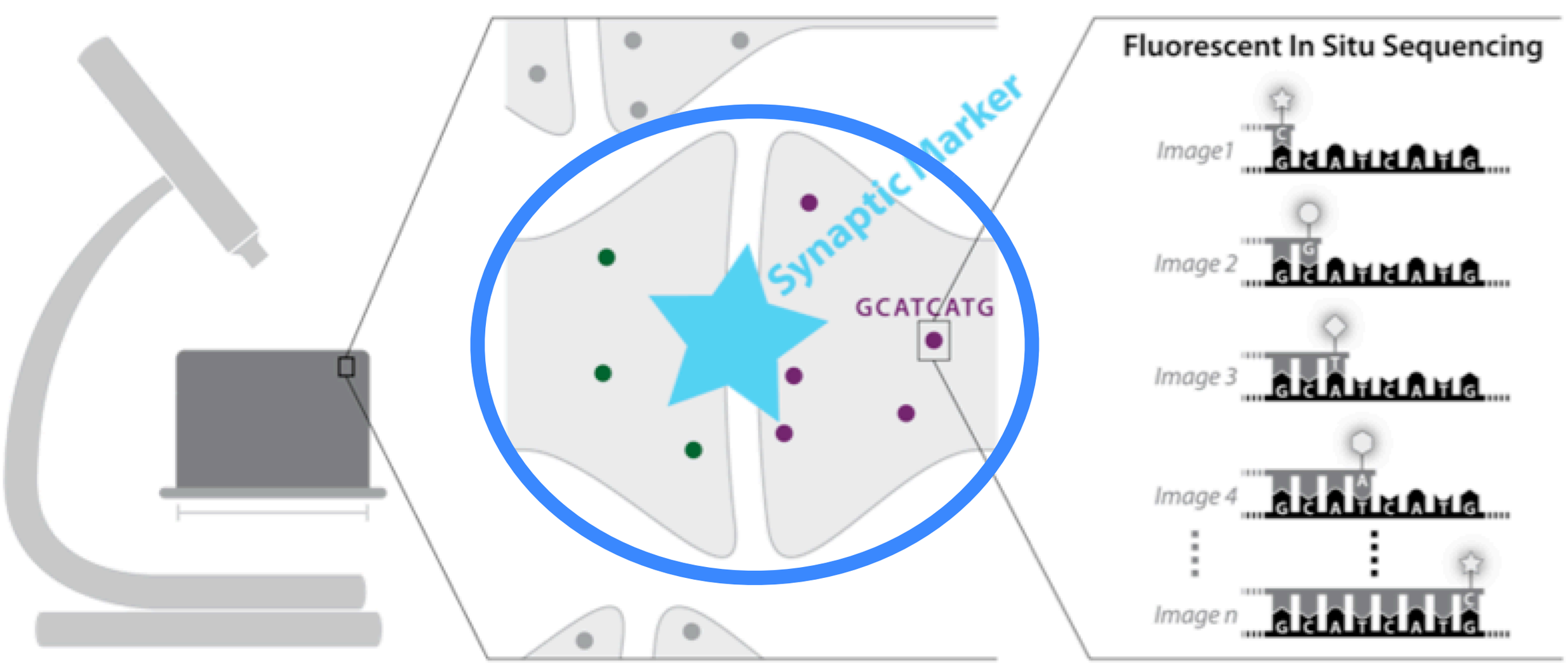
In-situ barcode connectomics

x,y,z ~ 600 nm

X,Y,Z > 1 cm

c ~ infinity

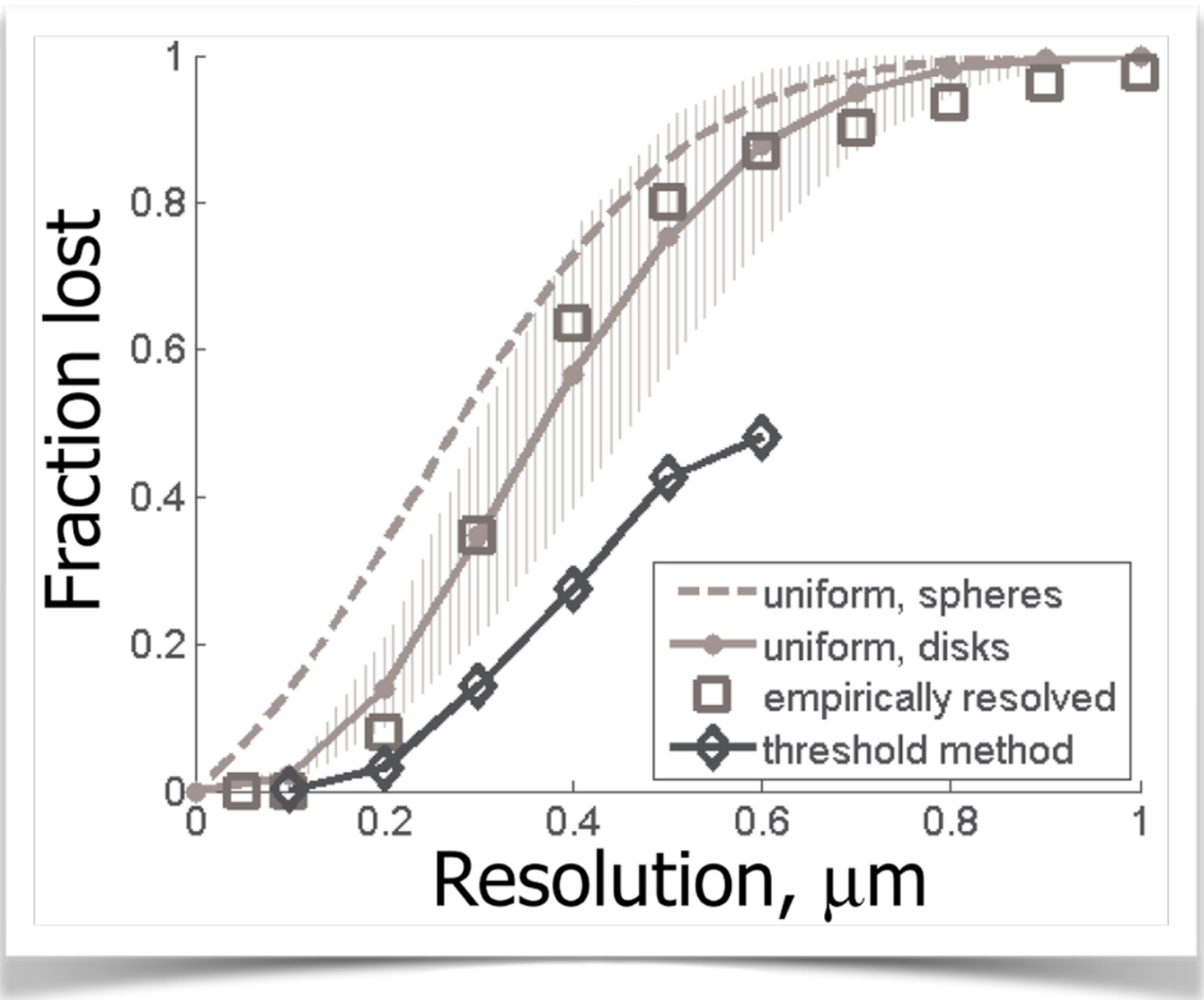
Synaptic resolution requires x,y,z < 100 nm



In-situ barcode connectomics

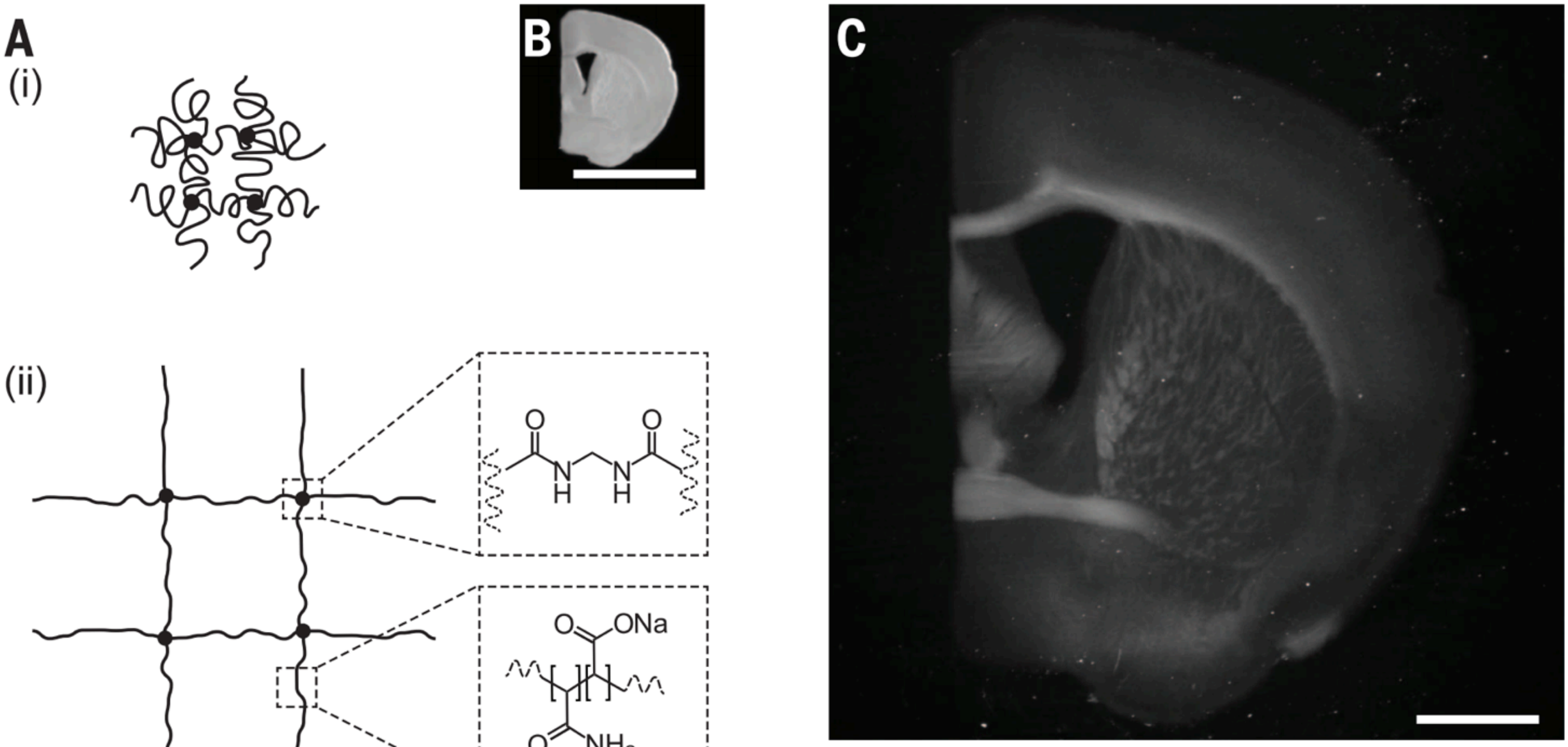
x,y,z ~ 600 nm
X,Y,Z > 1 cm
c ~ infinity

Synaptic resolution requires x,y,z < 100 nm



Synapse resolvability vs. spatial resolution:
Yuriy Mishchenko, 2011

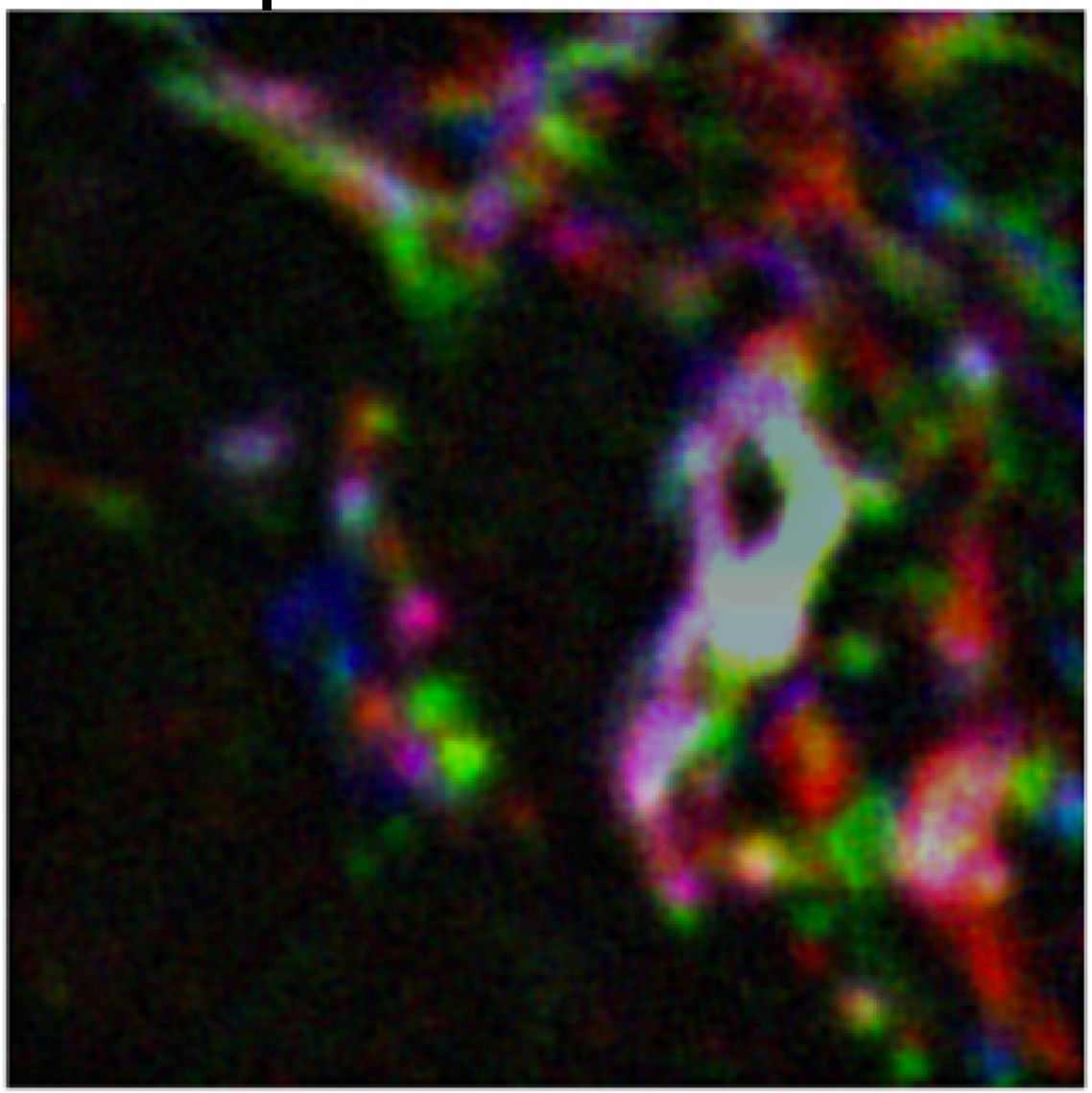
Expansion Microscopy (ExM)



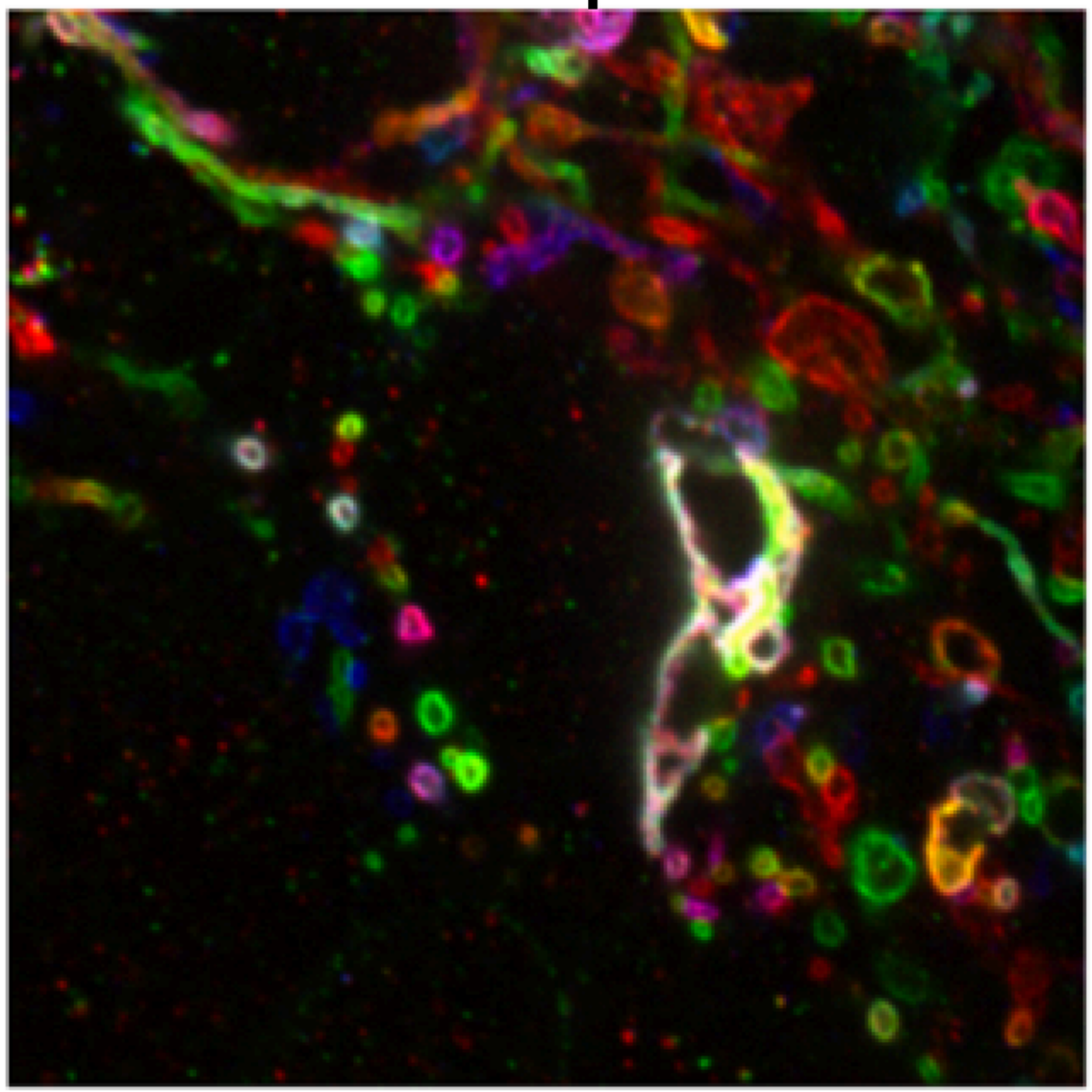
Chen*, Tillberg*, Boyden

Expansion Microscopy (ExM)

no expansion



4.5x linear expansion



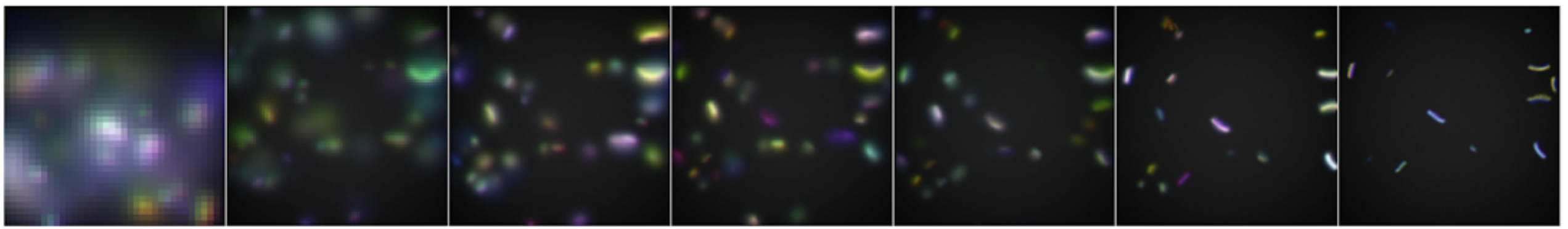
same microscope: a fast, diffraction-limited confocal

Boyden and Cai labs

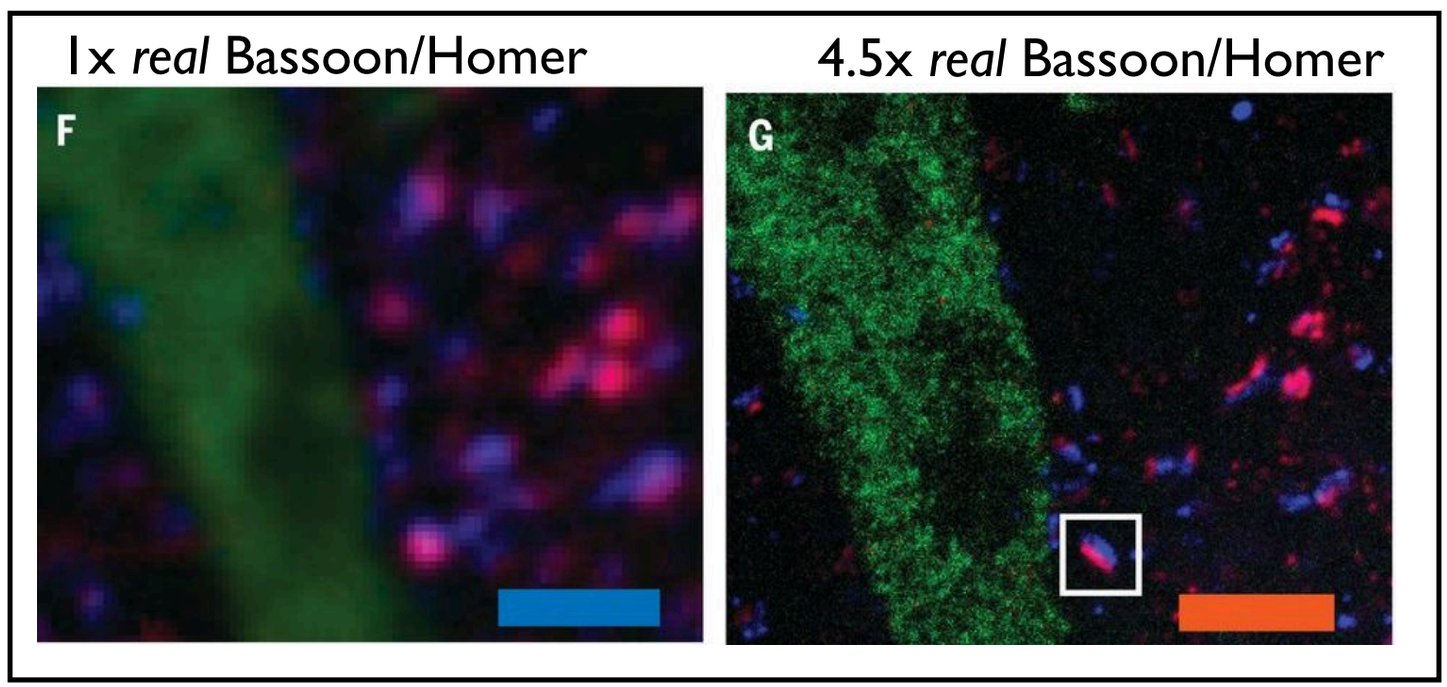
Expansion Microscopy (ExM)

Synapses at different levels of ExM expansion (simulated):

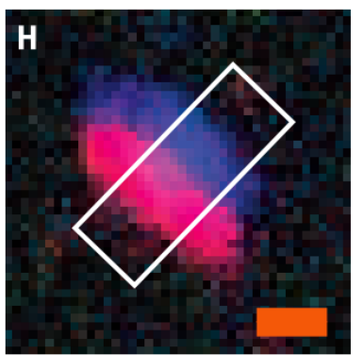
1x 2x 3x 4x 5x 10x 20x



Simulation based on Kasthuri et al published dataset
Color indicates cell of origin



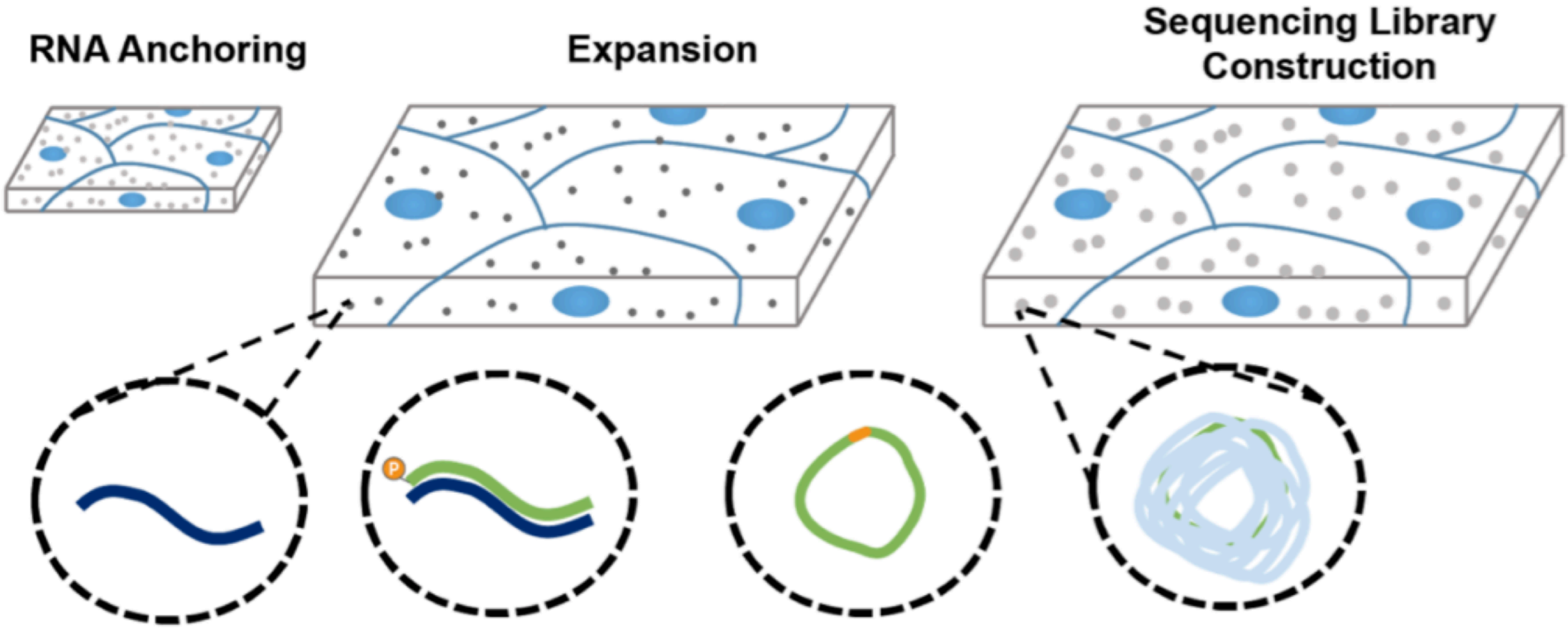
Pre/Post



4.5x ExM

Jeremy Wohlwend

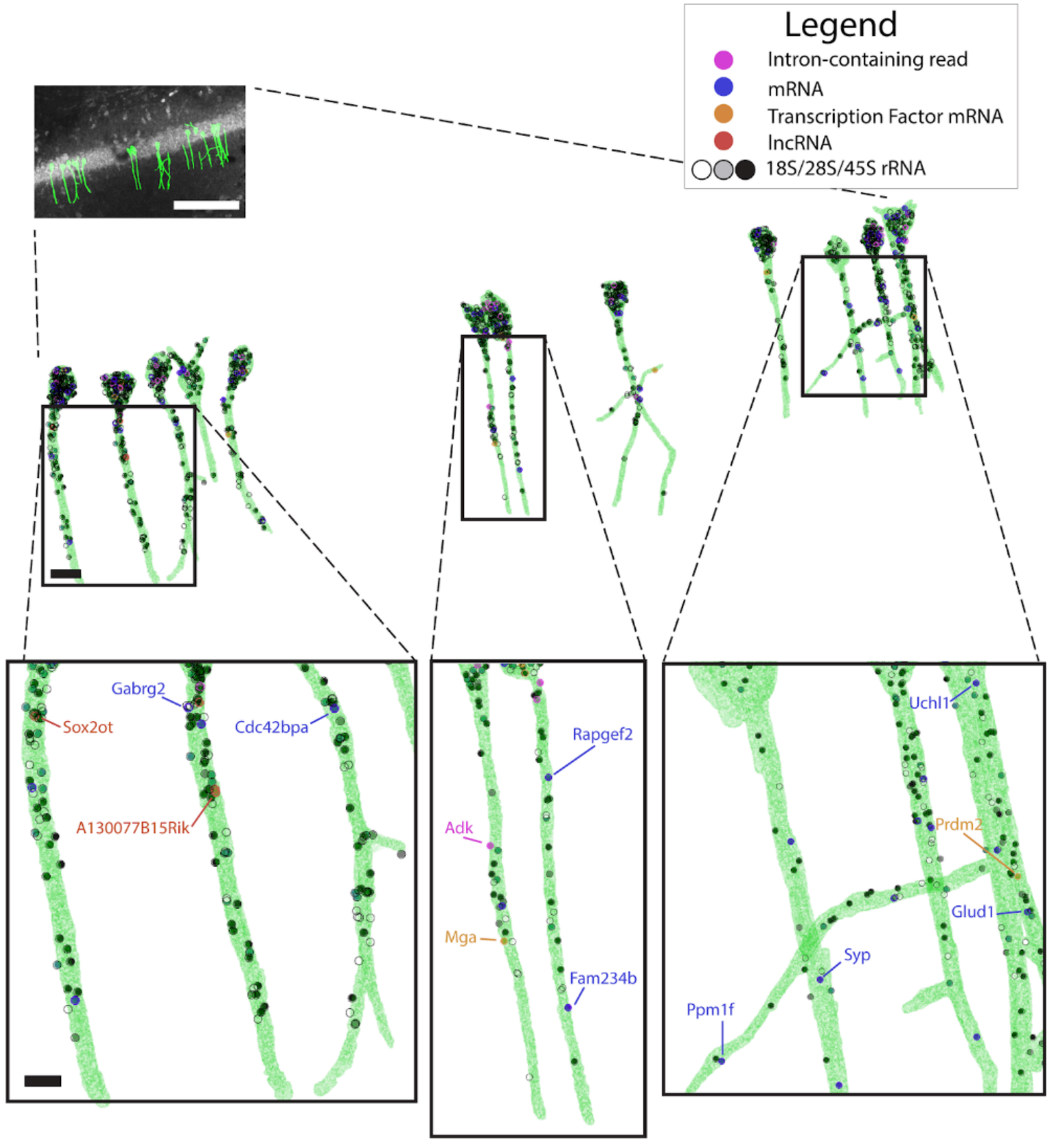
FISSEQ-ExM



Title: Expansion Sequencing: Spatially Precise *In Situ* Transcriptomics in Intact Biological Systems

Authors: Shahar Alon[1,2,3†], Daniel R Goodwin[1,2†], Anubhav Sinha[1,2,4†], Asmamaw T Wassie[1,2,5†], Fei Chen[1,6†], Evan R Daugharthy[7,8††], Yosuke Bando[1,9], Atsushi Kajita[10], Andrew G Xue[1], Karl Marrett[10], Robert Prior[10], Yi Cui[1,2], Andrew C Payne[1,6], Chun-Chen Yao[1,6], Ho-Jun Suk[1,2,4], Ru Wang[1,2], Chih-Chieh (Jay) Yu[1,2,5], Paul Tillberg[1†††], Paul Reginato[1,5,6,7,8], Nikita Pak[1,2,11], Songlei Liu[7,8], Sukanya Punthambaker[7,8], Eswar P. R. Iyer[8], Richie E Kohman[7,8], Jeremy A Miller[12], Ed S Lein[12], Ana Lako[13], Nicole Cullen[13], Scott Rodig[13], Karla Helvie[14], Daniel L Abravanel[14], Nikhil Wagle[14], Bruce E Johnson[14], Johanna Klughammer[6], Michal Slyper[6], Julia Waldman[6], Judit Jané-Valbuena[6], Orit Rozenblatt-Rosen[6], Aviv Regev[6,15,16], IMAXT Consortium[17], George M Church[7,8#*], Adam H Marblestone[1#†], Edward S Boyden[1,2,5,15,18#*]

FISSEQ-ExM

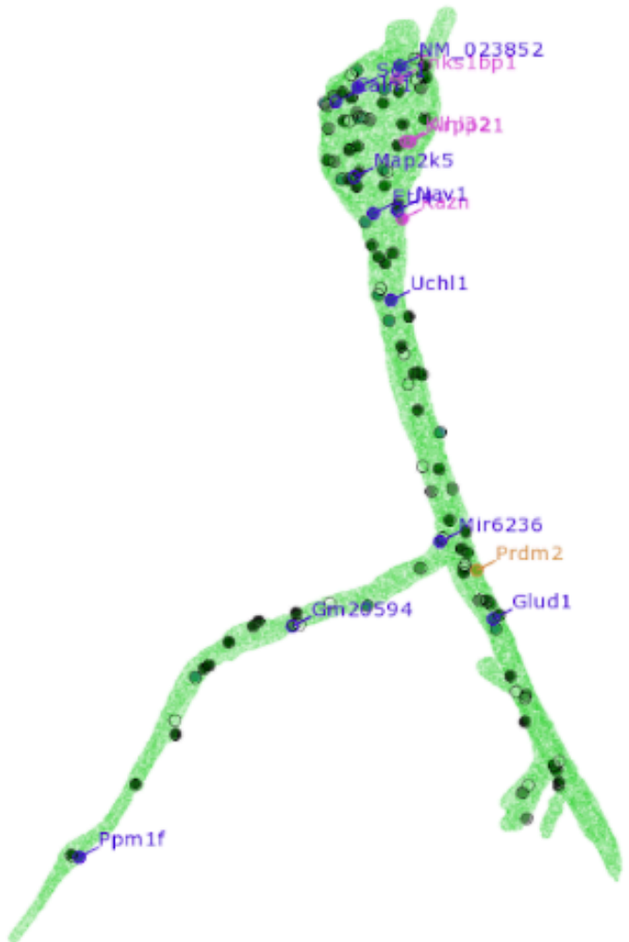


Alon*, Goodman*, Chen*, Daugharthy, …, Church**, Marblestone**, Boyden**.

FISSEQ-ExM

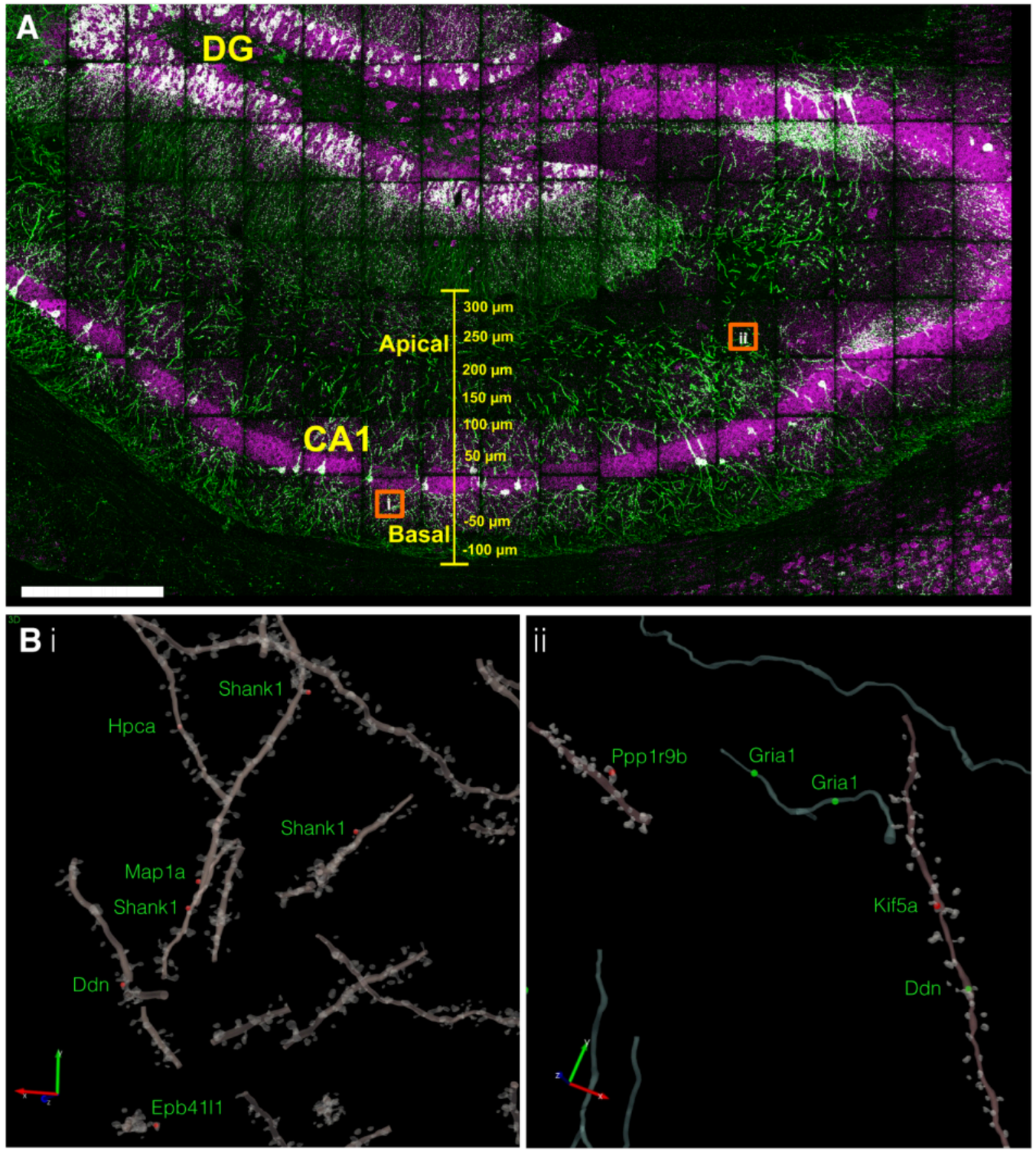
Neuron 13:

Gene Symbol	Gene name	Notes
Tnks1bp1	Tankyrase 1 Binding Protein 1	Intron
NM_023852	RAB3C, Member RAS Oncogene Family	
Kazn	Kazrin, Periplakin Interacting Protein	Intron
Etl4	Epilepsy, Occipitotemporal Lobe, And Migraine With Aura	
Caln1	Calneuron 1	
Nav1	Neuron Navigator 1	
Sos1	SOS Ras/Rac Guanine Nucleotide Exchange Factor 1	
Klhl32	Kelch Like Family Member 32	Intron
Arpp21	CAMP Regulated Phosphoprotein 21	Intron
Map2k5	Mitogen-Activated Protein Kinase Kinase 5	
Uchl1	Ubiquitin C-Terminal Hydrolase L1	
Mir6236		Probably rRNA
Prdm2	PR/SET Domain 2	Transcription Factor
Gm20594		Predicted gene
Glud1	Glutamate Dehydrogenase 1	
Ppm1f	Protein Phosphatase, Mg2+/Mn2+ Dependent 1F	



Alon*, Goodman*, Chen*, Daugharthy, …, Church**, Marblestone**, Boyden**.

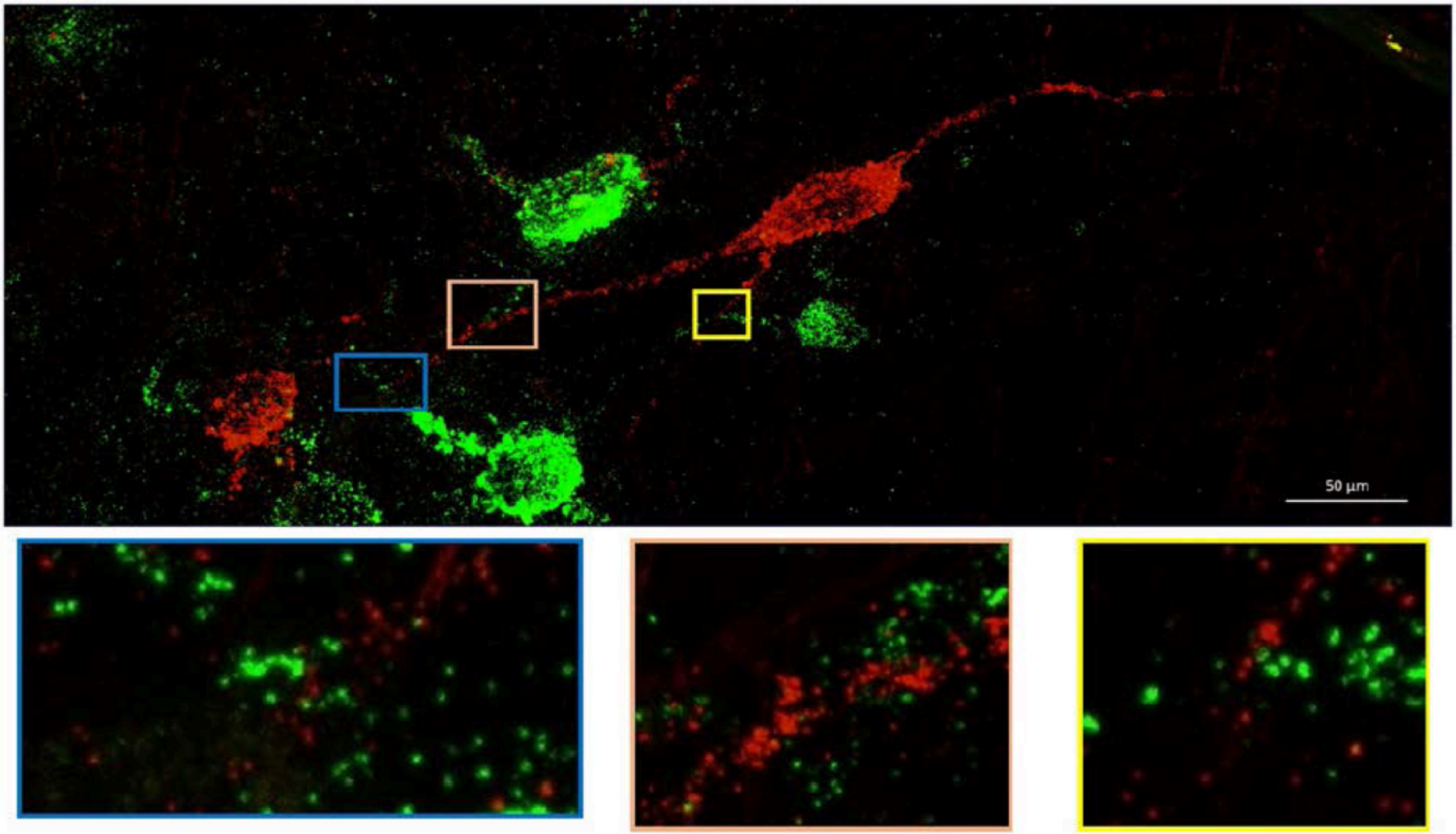
FISSEQ-ExM



Alon*, Goodman*, Chen*, Daugharthy, …, Church**, Marblestone**, Boyden**.

Scalable brain mapping: unique ID's for structure

Reading neuron barcode RNA sequences
*in **physically expanded** brain tissue*

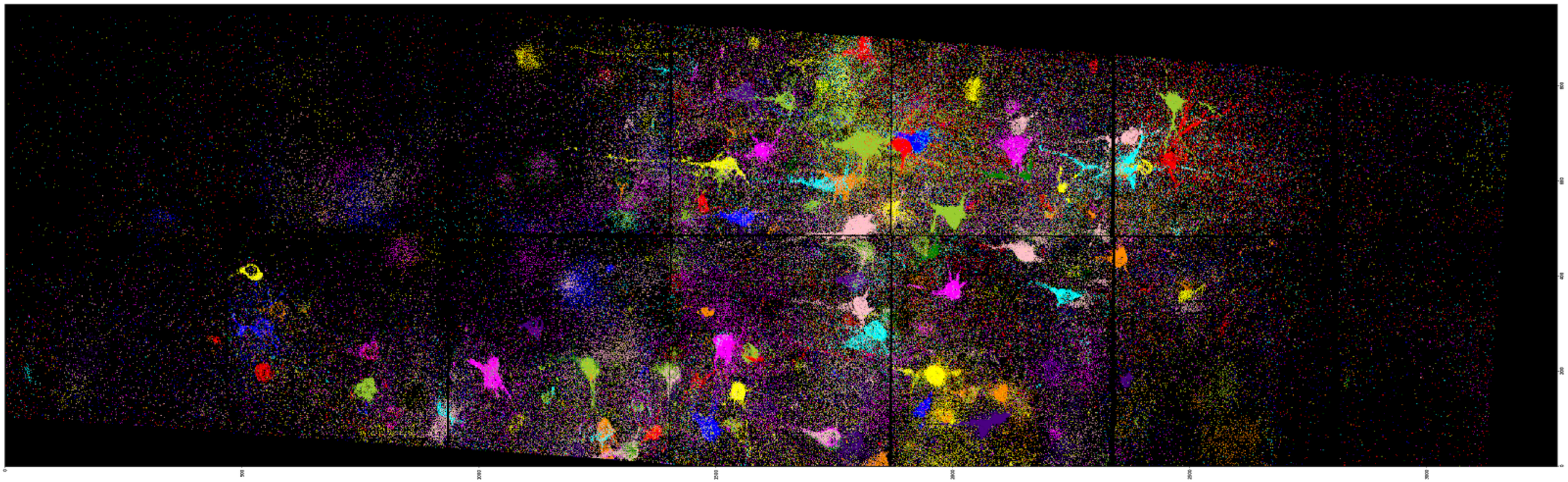


Data by Richie Kohman, analysis by Andrew Xue, Dan Goodwin, Ruihan Zhang

Scalable brain mapping: unique ID's for structure

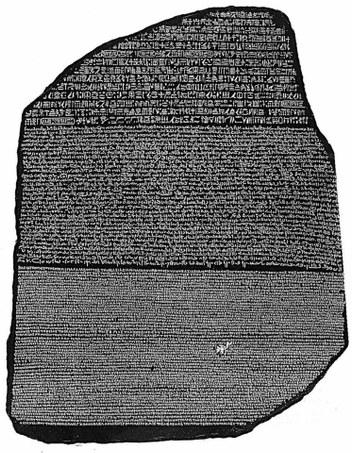
Reading neuron barcode RNA sequences
*in **physically expanded** brain tissue*

full thickness of mouse cortex



(first-pass *automatic* barcode extraction)

Data by Richie Kohman, analysis by Andrew Xue, Dan Goodwin, Ruihan Zhang et al



“Rosetta Brain”

Activity

Behavior

Connectivity

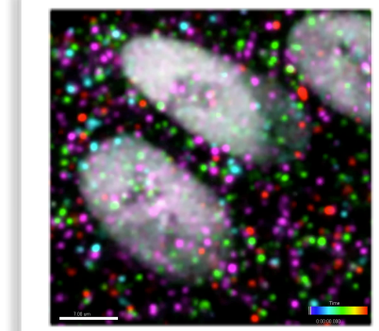
Development

Expression

? Molecular Tickertapes ?

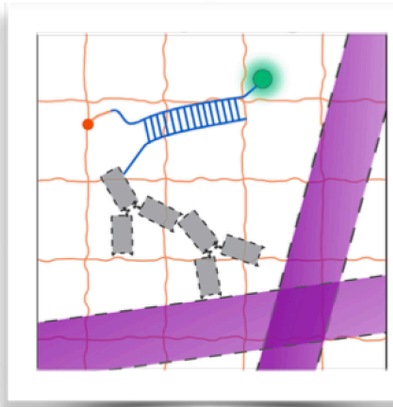
(in-vivo-generated)
Cell **Barcodes**
(update 1x per division)

RNA Transcripts +
DNA-barcoded **Antibodies**



FISSEQ
(Lee et al 2014)

+



ExM
(Chen et al 2014)

Rapidly evolving homing CRISPR barcodes

Reza Kalhor, Prashant Mali & George M Church

Large-scale simultaneous measurement of epitopes and transcriptomes in single cells.

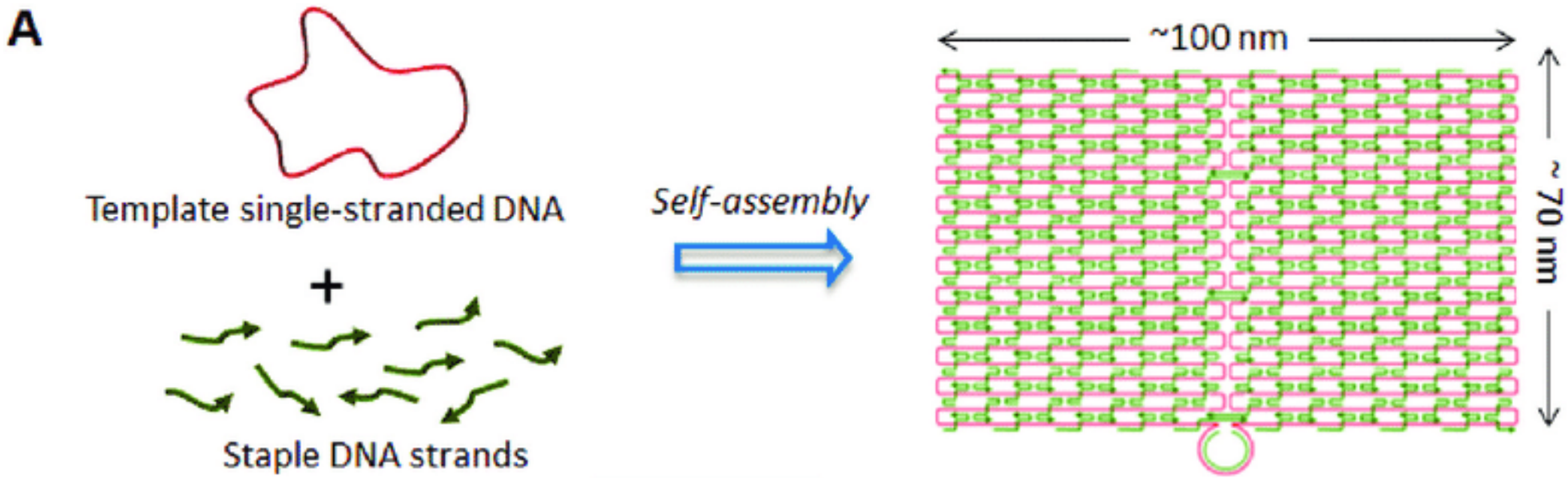
Marlon Stoeckius, Christoph Hafemeister, William Stephenson, Brian Houck-Loomis, Harold Swerdlow, Rahul Satija, Peter Smibert

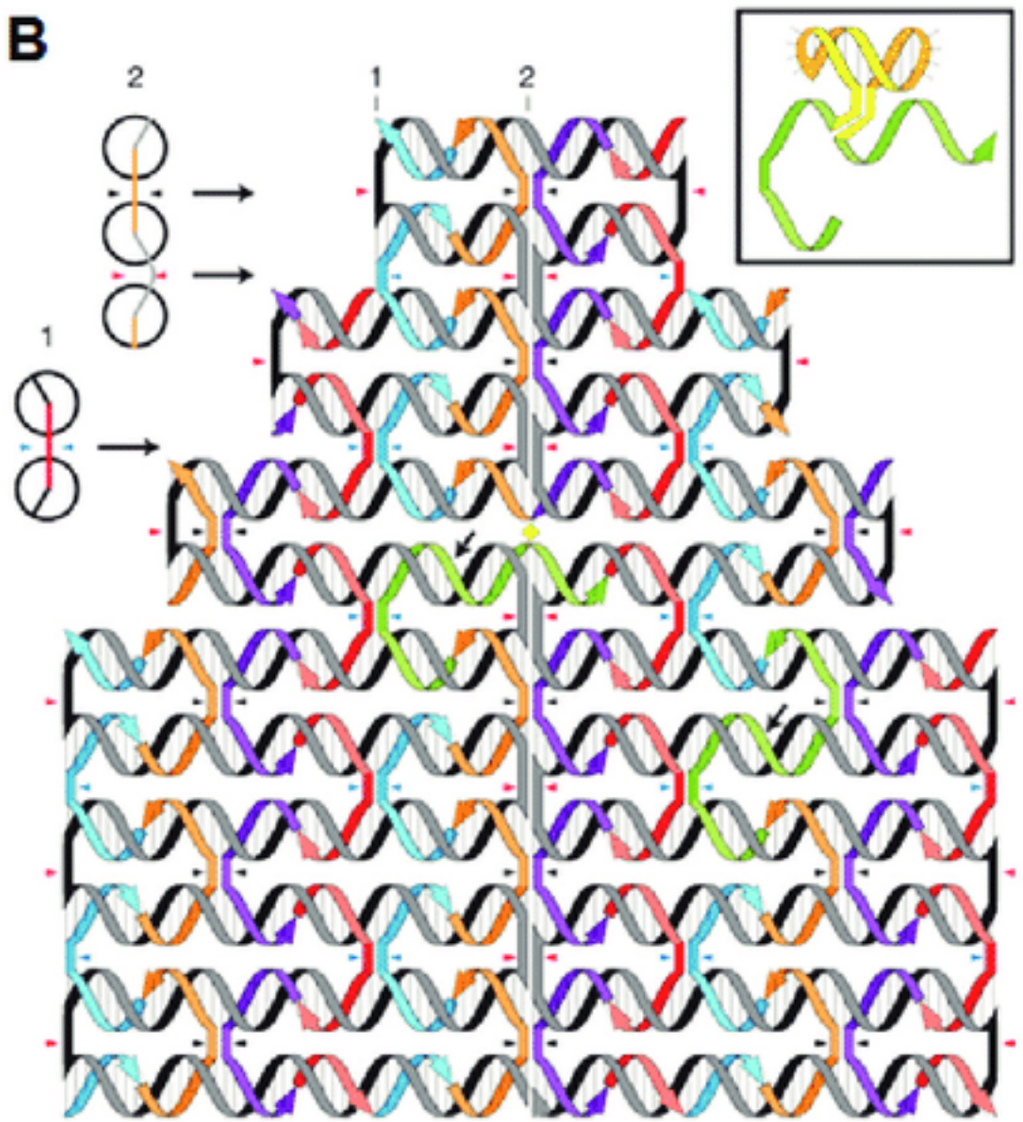
doi: https://doi.org/ 10.1101/113068

Part 2: construct

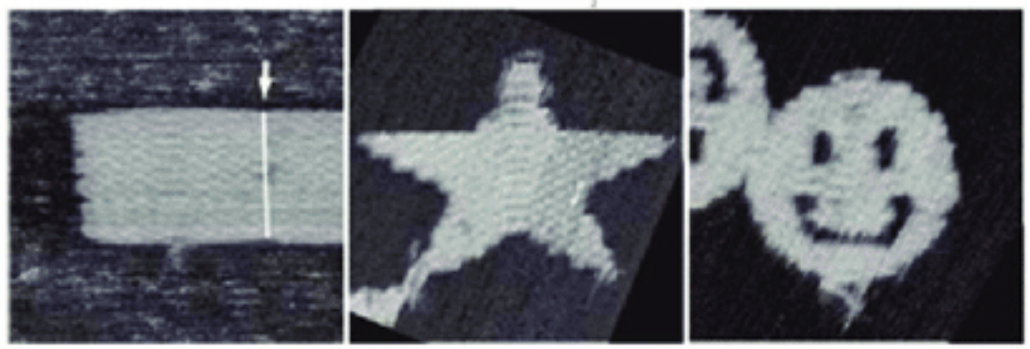
How to position molecules where we want

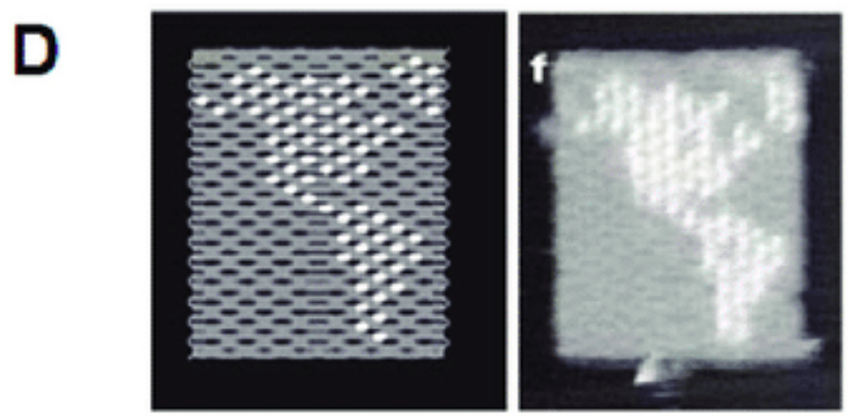
DNA origami


A
Template single-stranded DNA
+
Staple DNA strands
Self-assembly
~100 nm
~ 70 nm

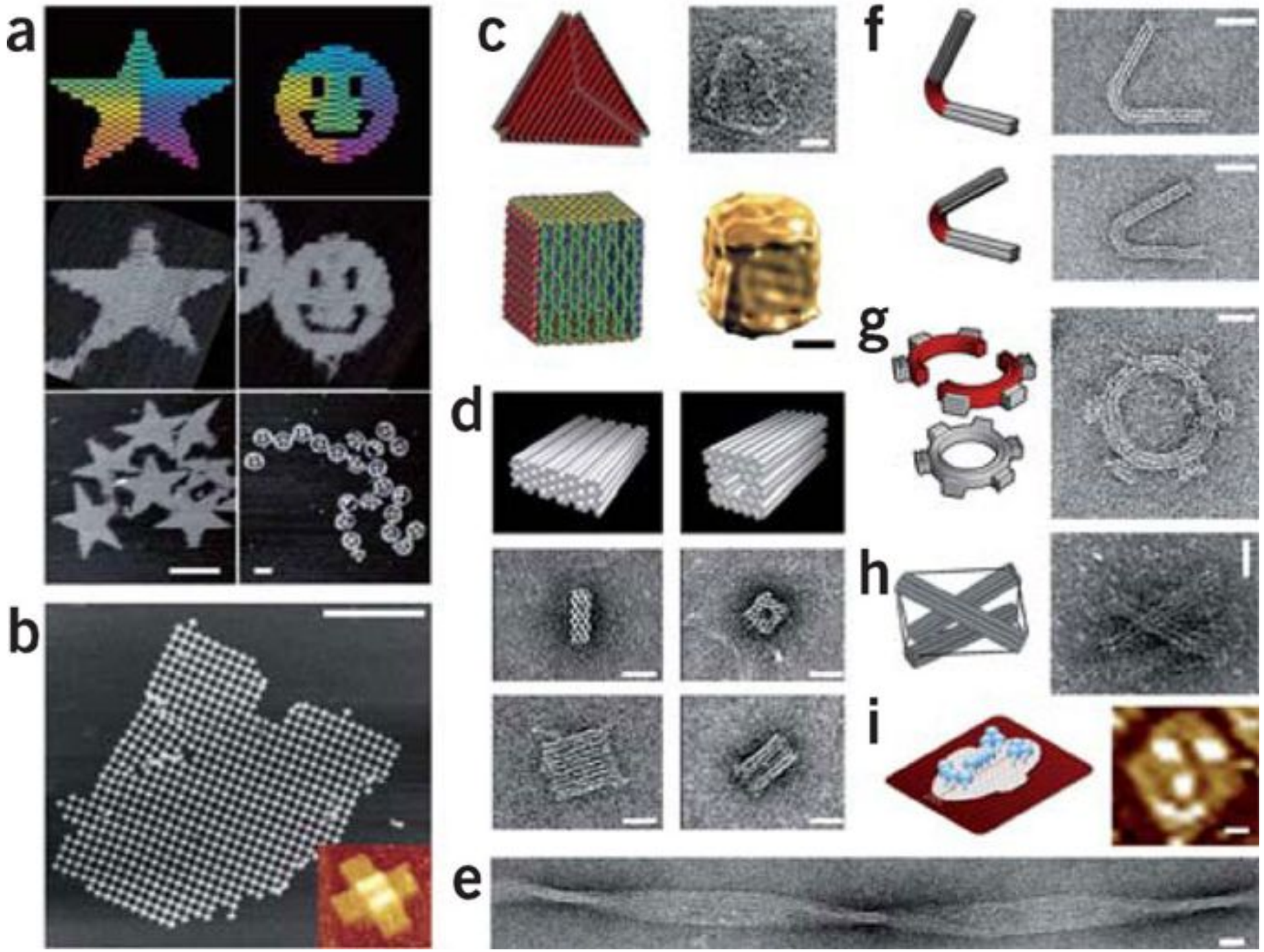

B
2
1
2
1


C
~ 100 nm



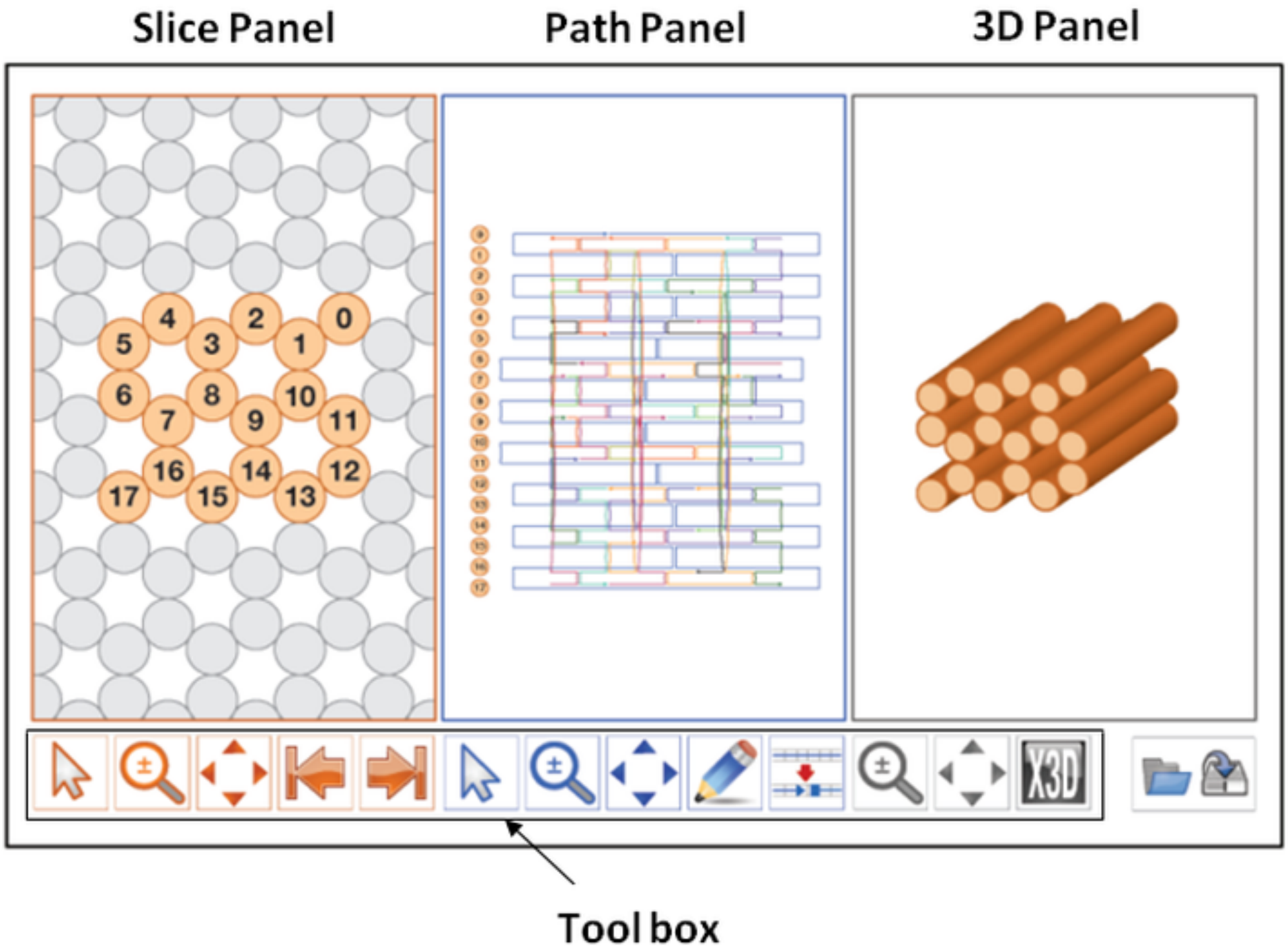

D
f

DNA origami



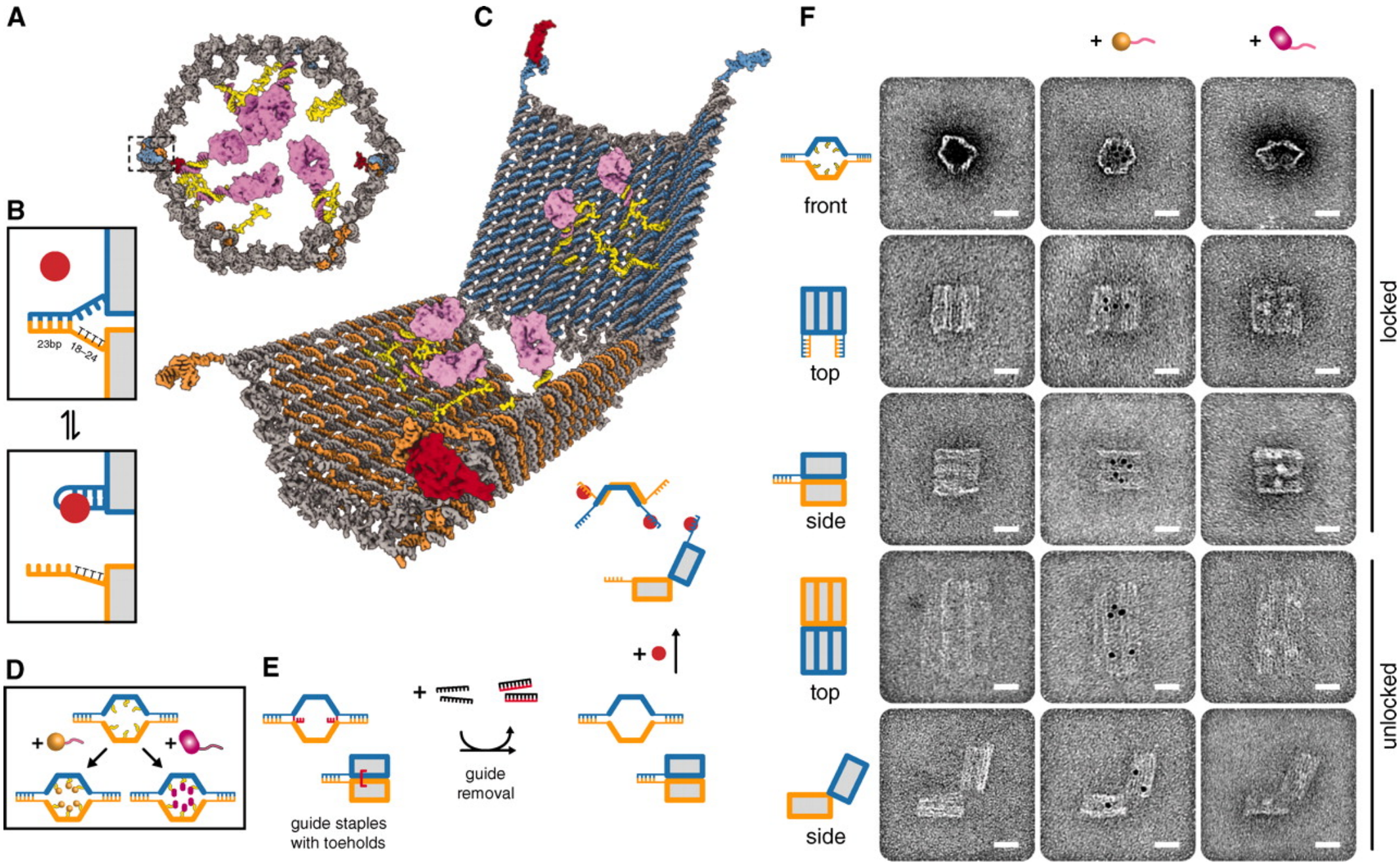
DNA origami

caDNAno: www.cadnano.org



Douglas, Marblestone et al, 2009

DNA origami



Douglas, Bachelet, 2012

DNA origami

x,y,z ~ 5-10 nm

DNA base spacing = 0.3 nm

helical diameter = 2 nm

orderly manner on solid substrates[11,12]. So far, few objects afford a design accuracy better than 5 nm[13–16], and the sub-nanometre scale has been reached only within the unit cells of designed DNA crystals[17]. Here, we report a molecular position-

X,Y,Z > 100 nm

c ~ no practical limit

NOT quite *small* enough for covalent "APM"

NOT quite *big* enough for integration into devices

DNA origami

x,y,z ~ 5-10 nm

DNA base spacing = 0.3 nm

helical diameter = 2 nm

orderly manner on solid substrates[11,12]. So far, few objects afford a design accuracy better than 5 nm[13–16], and the sub-nanometre scale has been reached only within the unit cells of designed DNA crystals[17]. Here, we report a molecular position-

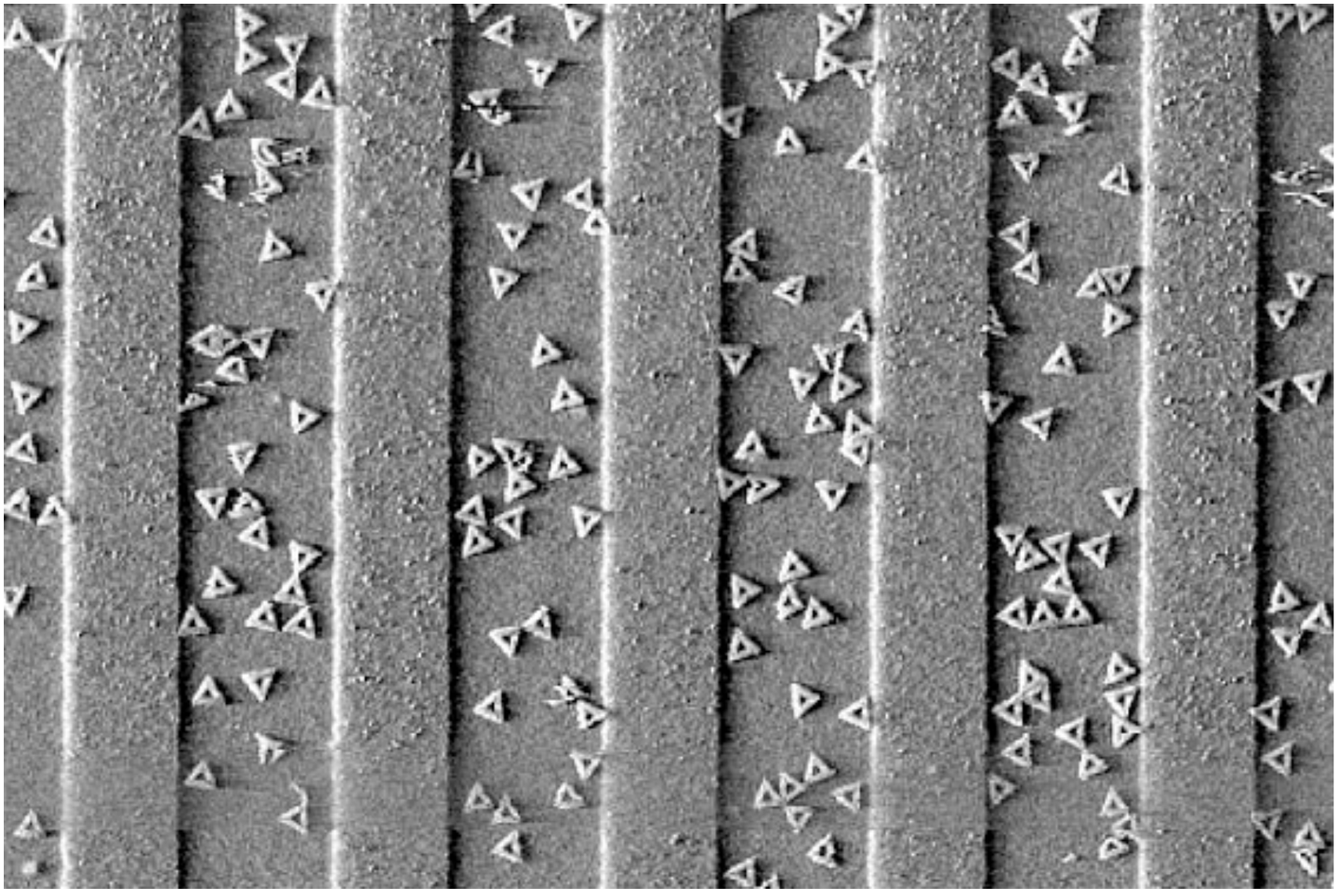
X,Y,Z > 100 nm

c ~ no practical limit

NOT quite *small* enough for covalent "APM"

NOT quite *big* enough for integration into devices

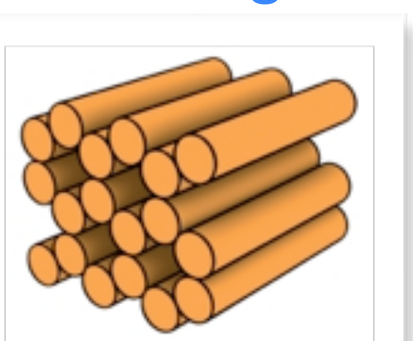
How to combine “top-down” with “bottom-up” in nanotechnology?



Rothemund 2009

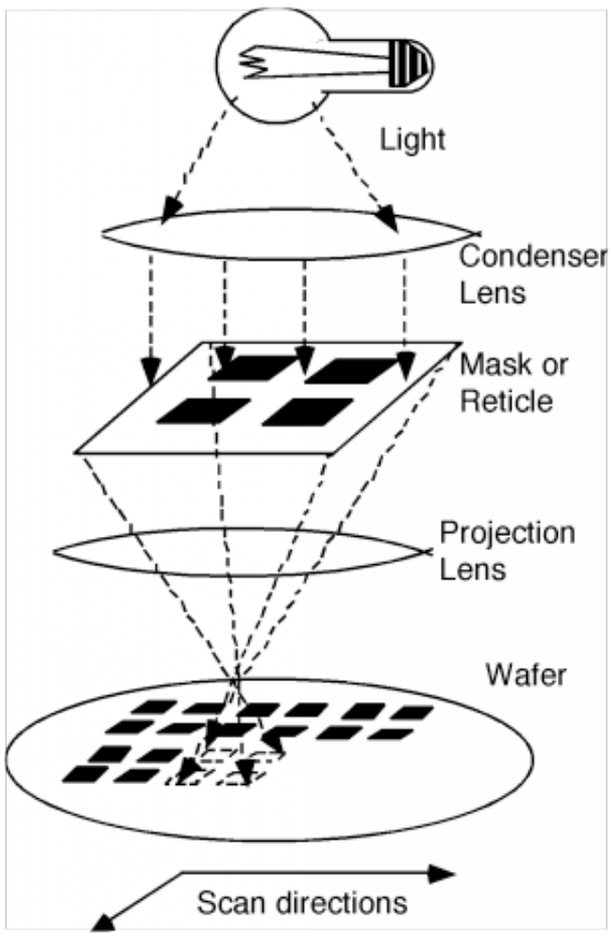
How to combine "top-down" with "bottom-up" in nanotechnology?

DNA origami



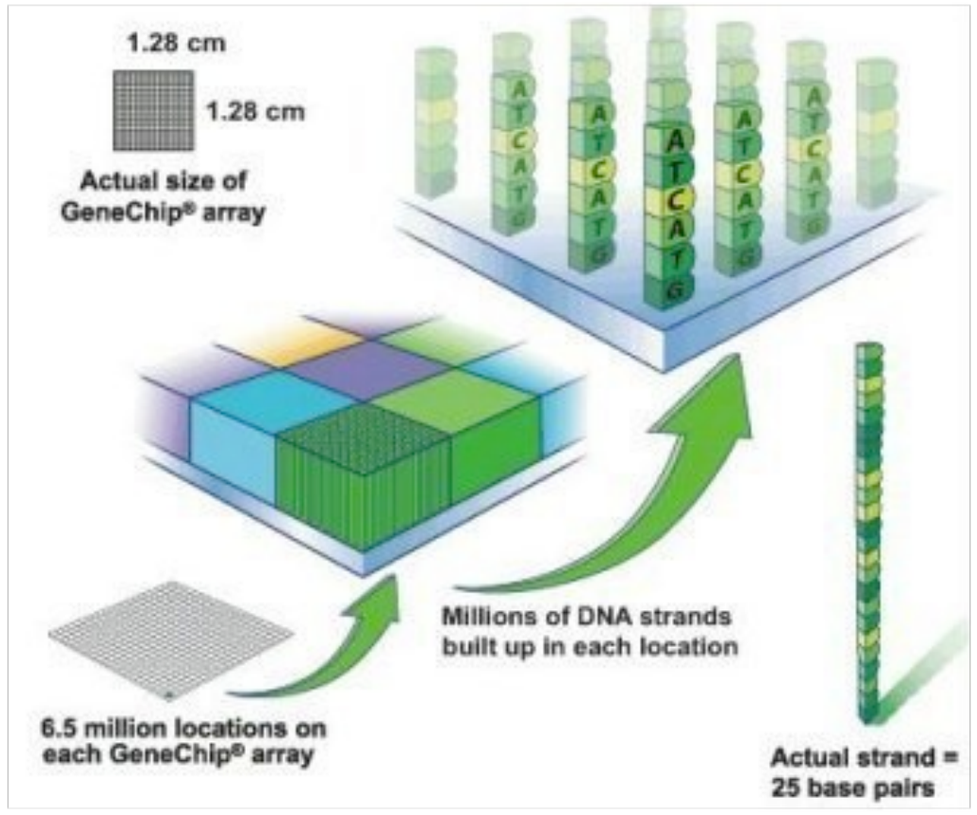
+

photolithography



+

DNA micro-arrays



= **chip-scale** programmable nano-fabrication

[Church, Marblestone et al]

How to combine "top-down" with "bottom-up" in nanotechnology?

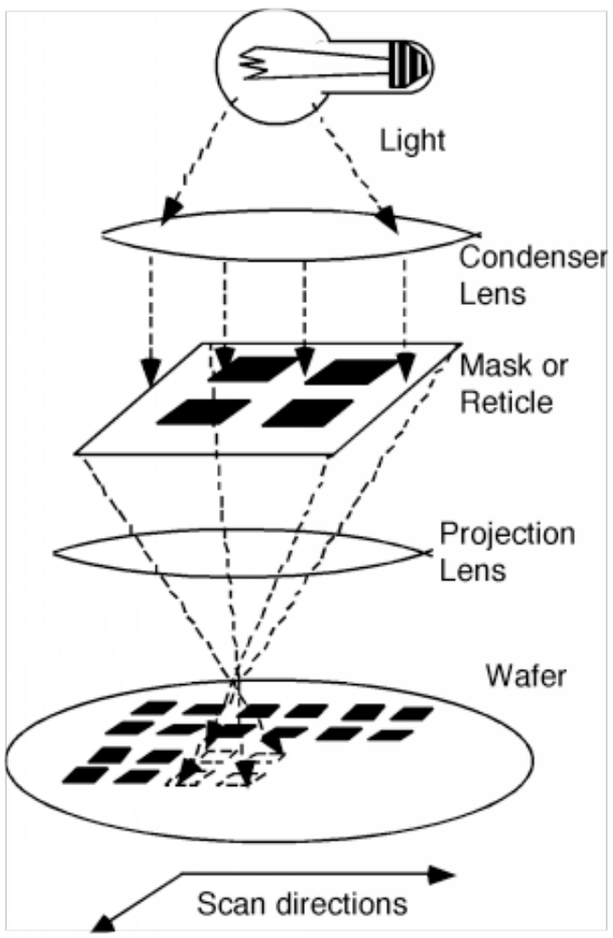
e.g., light-directed DNA synthesis on chip

↓

DNA origami

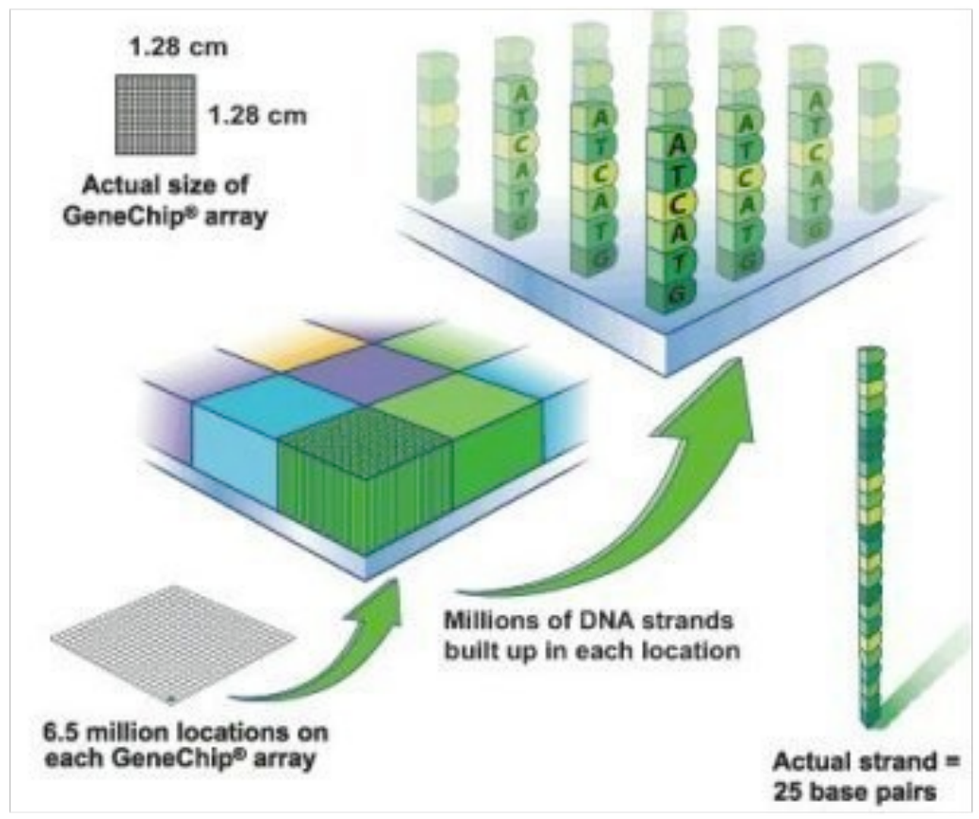
+

photolithography



+

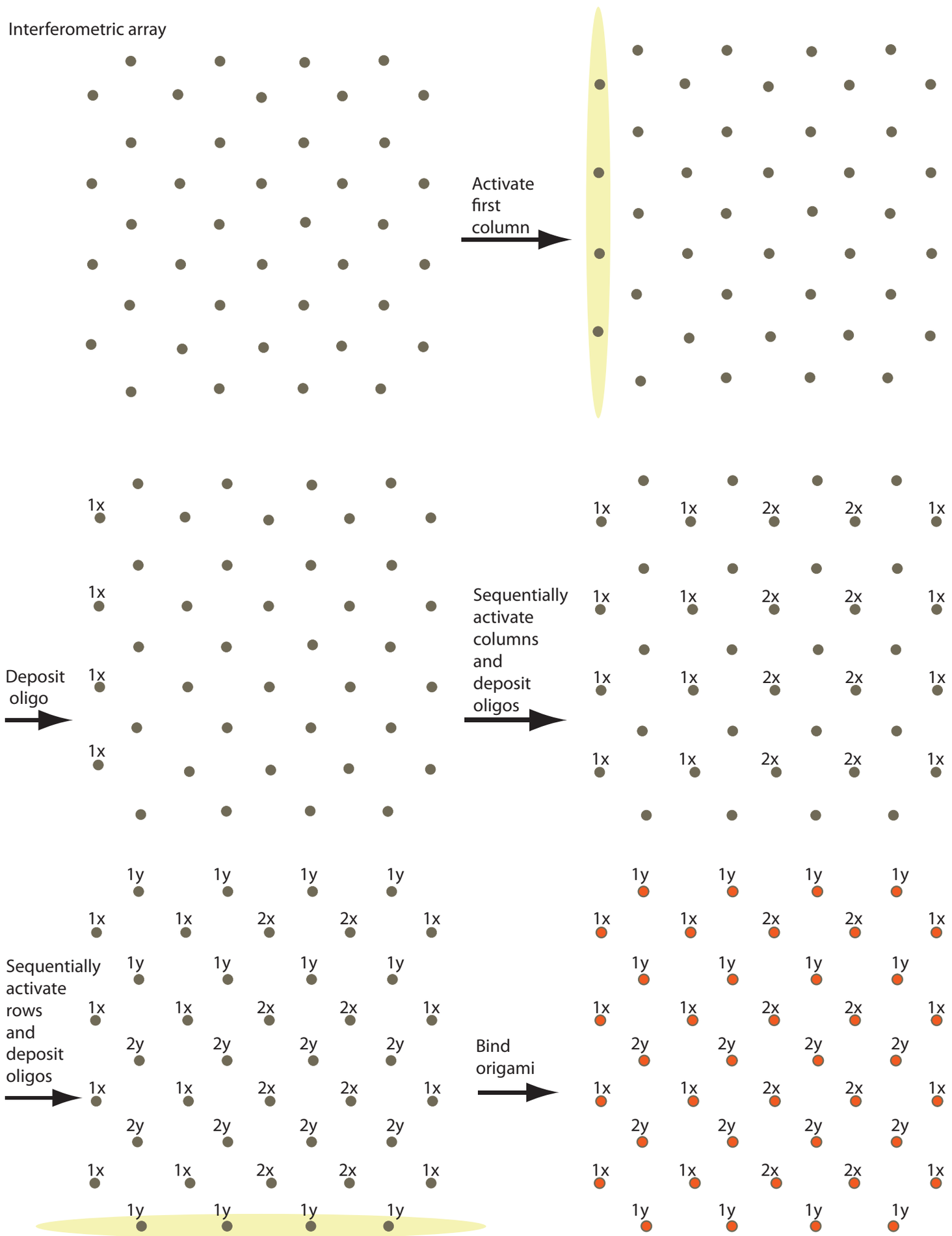
DNA micro-arrays



= **chip-scale** programmable nano-fabrication

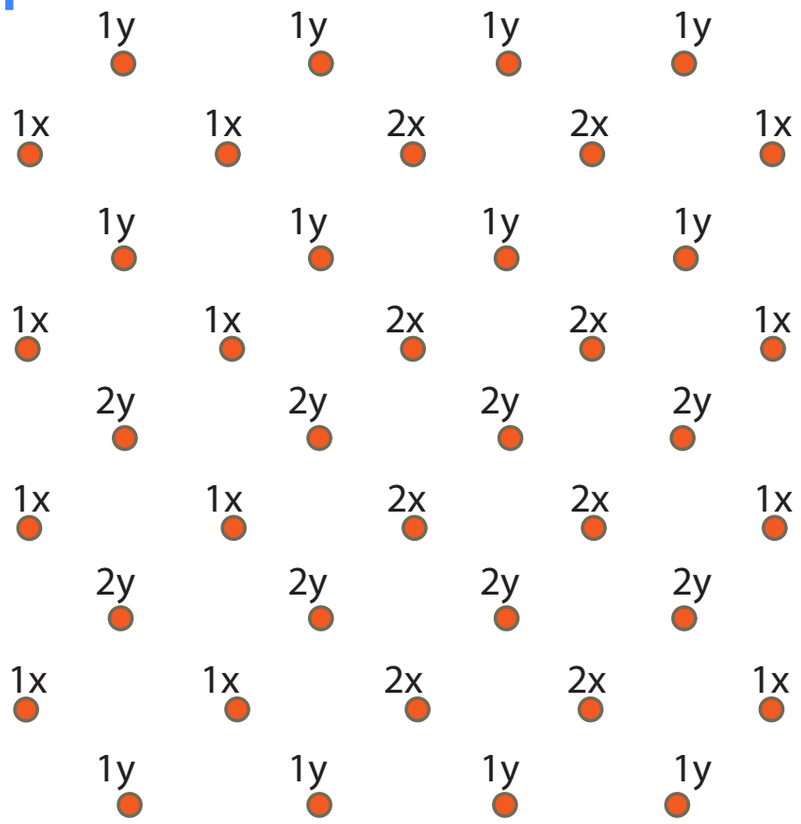
[Church, Marblestone et al]

early “nm2cm”

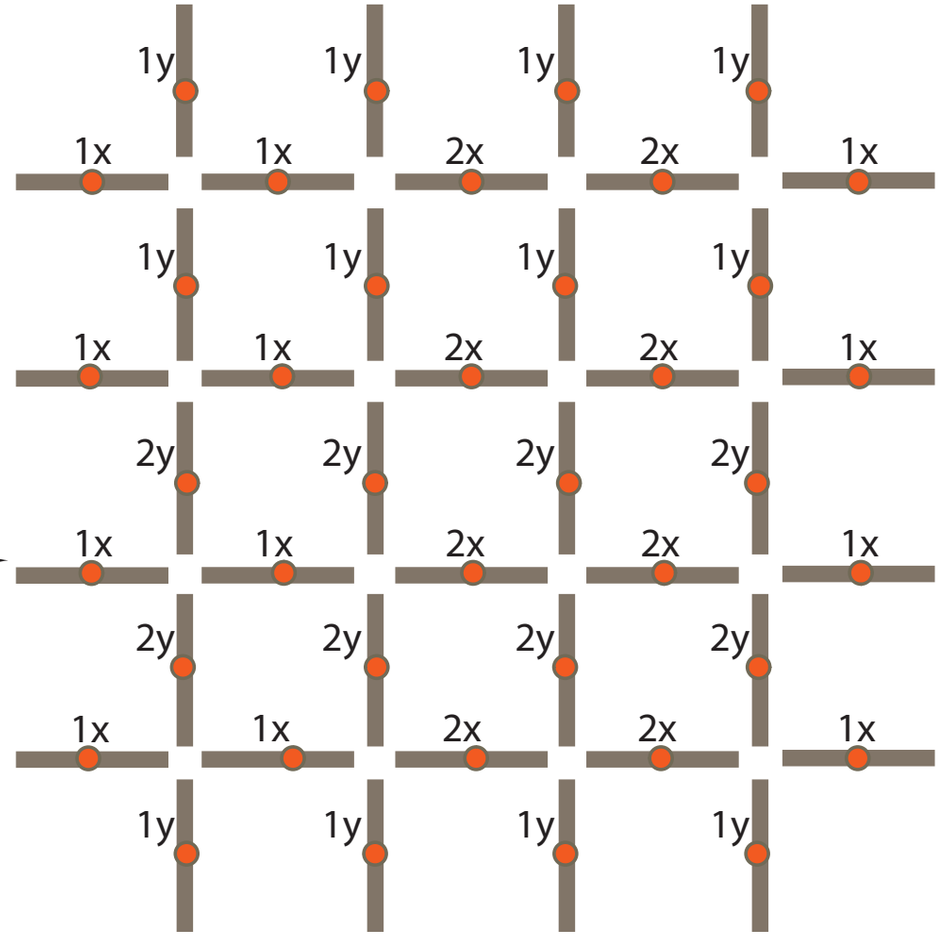

Interferometric array
Activate first column
1x
1x
Sequentially activate columns and deposit oligos
1x
1x
2x
2x
1x
Deposit oligo
1x
1x
Sequentially activate rows and deposit oligos
1y
1y
1y
1y
1x
1x
2x
2x
1x
2y
2y
2y
2y
Bind origami

early "nm2cm"

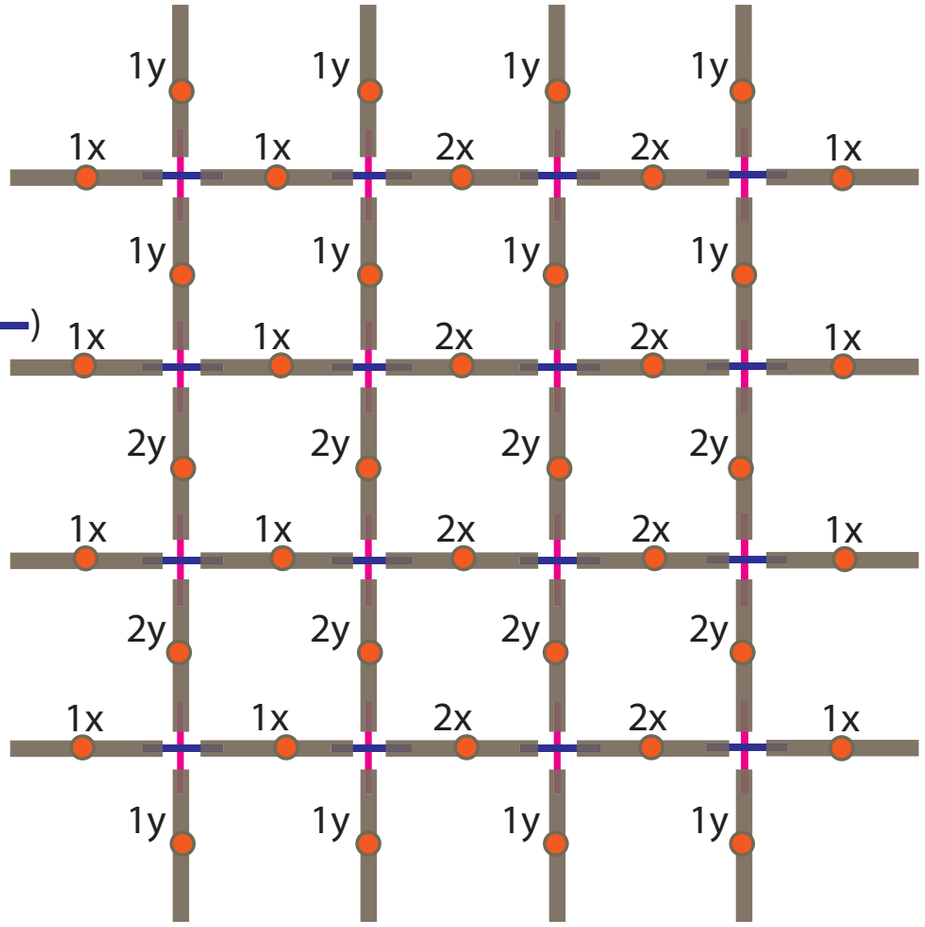
DNA
nano-array
synthesis



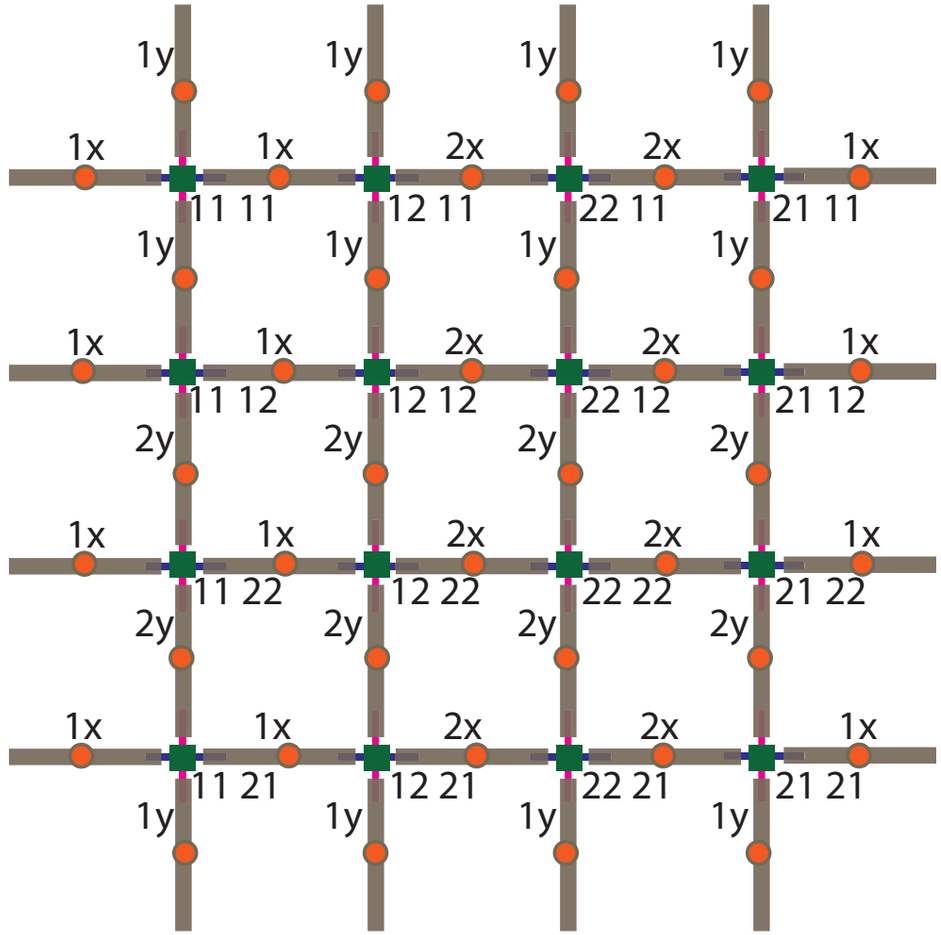
Bind DNA
nano-rods



Add
vertical (▮)
and
horizontal (▬)
rod
coupling
DNAs
(RCDs)



Add
2nd order
rod
coupling
DNAs (■)
(to couple
RCD pairs)

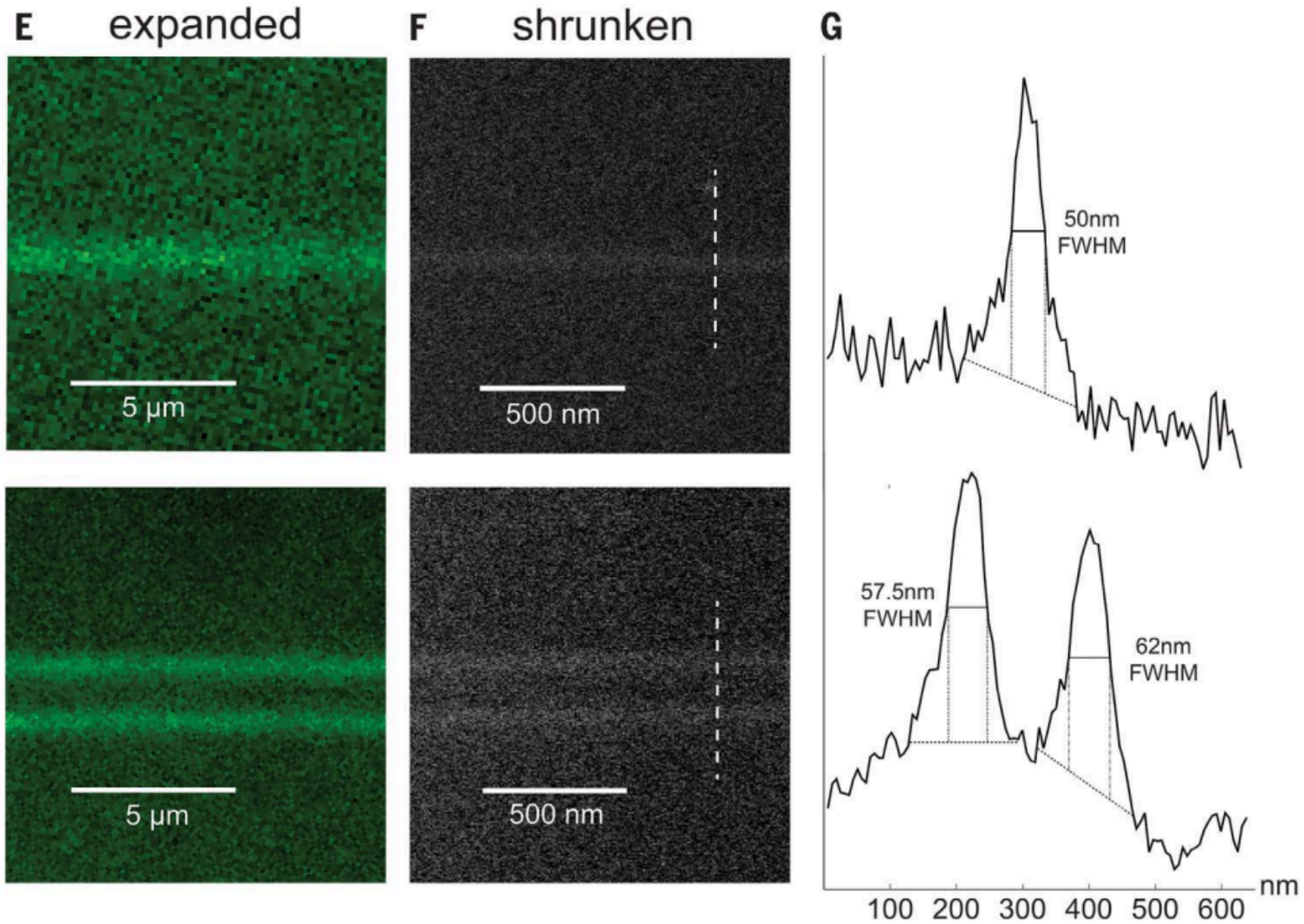


How to combine “top-down” with “bottom-up” in nanotechnology?

3D nanofabrication by volumetric deposition and controlled shrinkage of patterned scaffolds

Daniel Oran[1,*], Samuel G. Rodriques[1,2,*], Ruixuan Gao[1], Shoh Asano[1,3], Mark A. Skylar-Scott[4,5], Fei Chen[1,6], Paul W. Tillberg[1,7,†], Adam H. Marblestone[1,‡], Edward S. Boyden[1,6,8,9,10,‡,§]

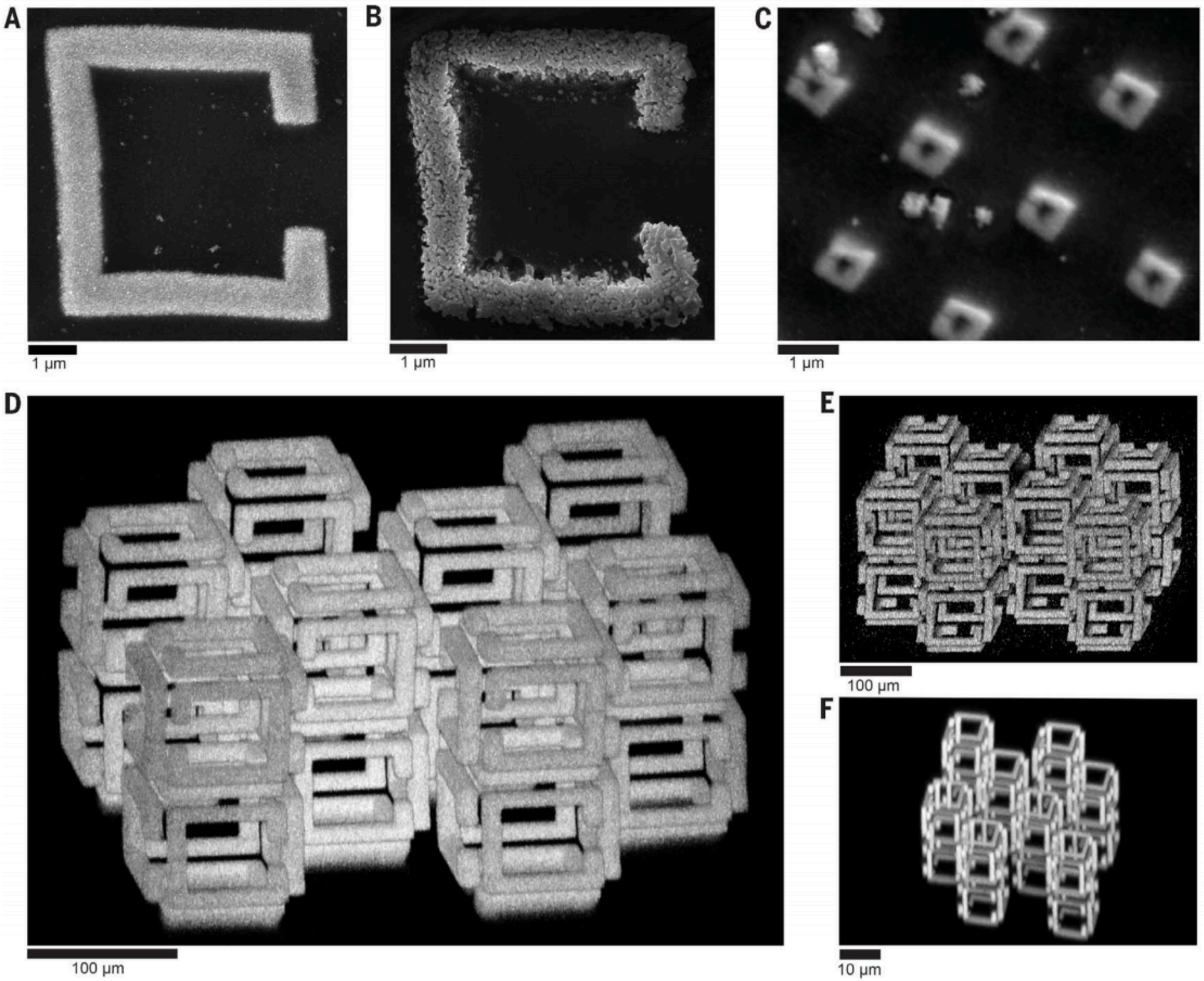
How to combine "top-down" with "bottom-up" in nanotechnology?



3D nanofabrication by volumetric deposition and controlled shrinkage of patterned scaffolds

Daniel Oran[1,*], Samuel G. Rodriques[1,2,*], Ruixuan Gao[1], Shoh Asano[1,3], Mark A. Skylar-Scott[4,5], Fei Chen[1,6], Paul W. Tillberg[1,7,†], Adam H. Marblestone[1,‡], Edward S. Boyden[1,6,8,9,10,‡,§]

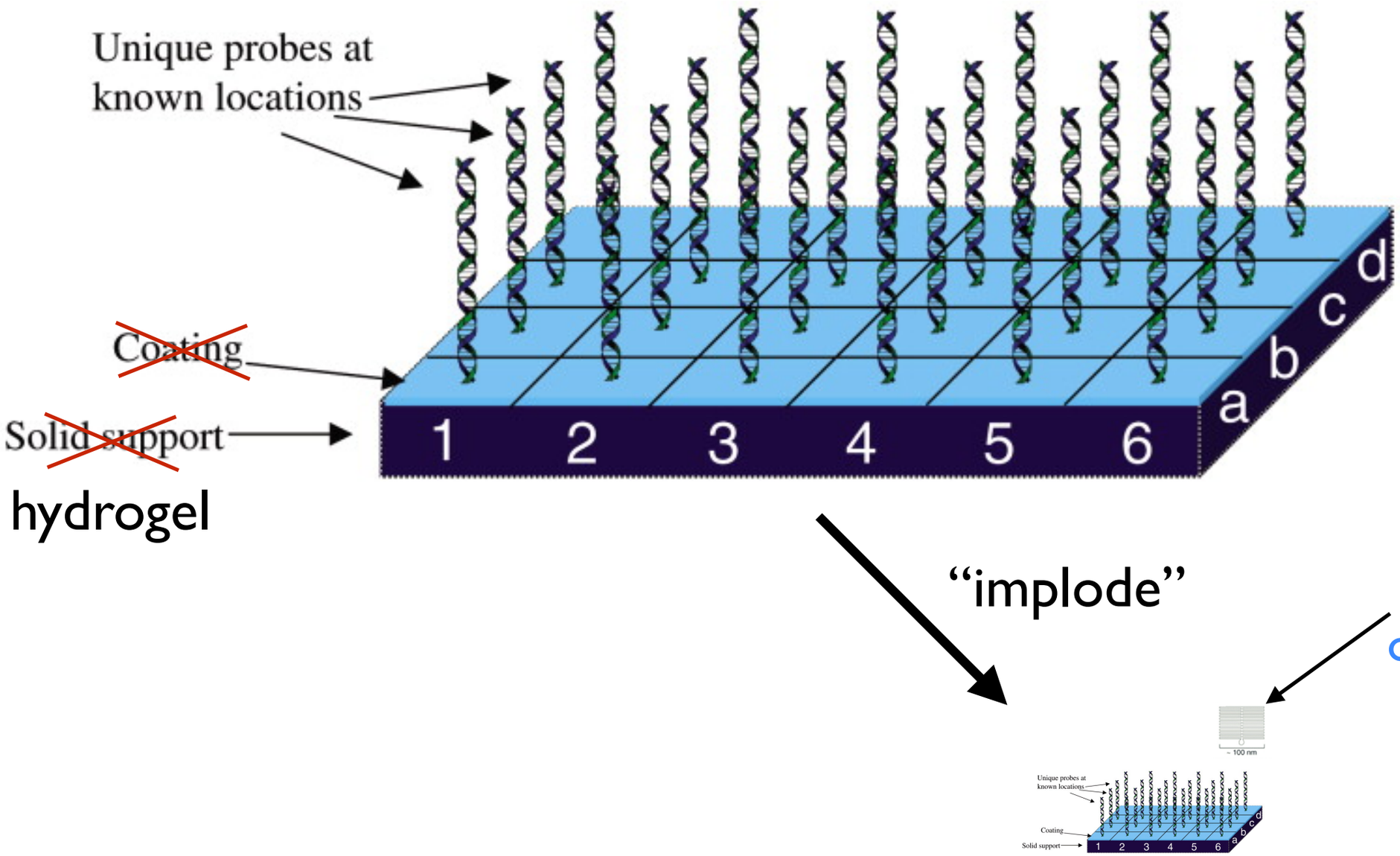
How to combine "top-down" with "bottom-up" in nanotechnology?



3D nanofabrication by volumetric deposition and controlled shrinkage of patterned scaffolds

Daniel Oran[1,*], Samuel G. Rodriques[1,2,*], Ruixuan Gao[1], Shoh Asano[1,3], Mark A. Skylar-Scott[4,5], Fei Chen[1,6], Paul W. Tillberg[1,7,†], Adam H. Marblestone[1,‡], Edward S. Boyden[1,6,8,9,10,‡,§]

How to combine "top-down" with "bottom-up" in nanotechnology?



DNA origami

x,y,z ~ 5-10 nm

DNA base spacing = 0.3 nm

helical diameter = 2 nm

orderly manner on solid substrates[11,12]. So far, few objects afford a design accuracy better than 5 nm[13–16], and the sub-nanometre scale has been reached only within the unit cells of designed DNA crystals[17]. Here, we report a molecular position-

X,Y,Z > 100 nm

c ~ no practical limit

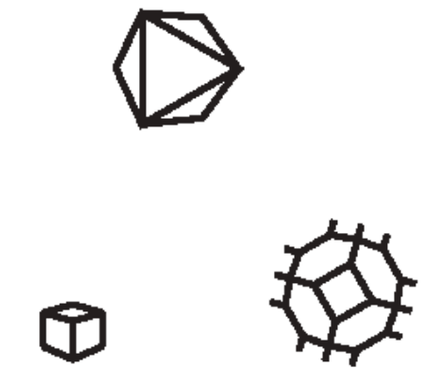
NOT quite *small* enough for covalent "APM"

NOT quite *big* enough for integration into devices

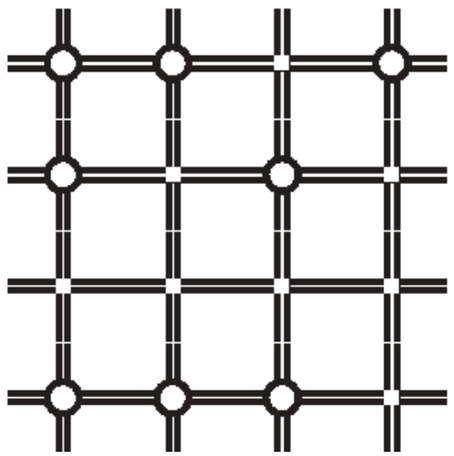
Going smaller…

previous work

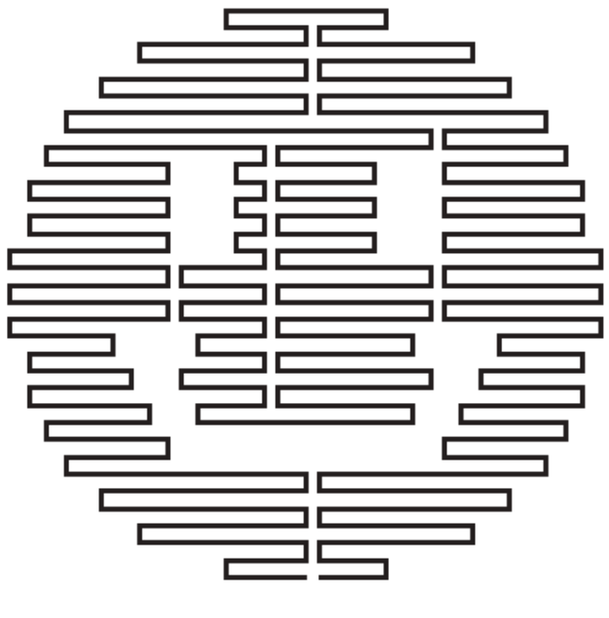
shapes

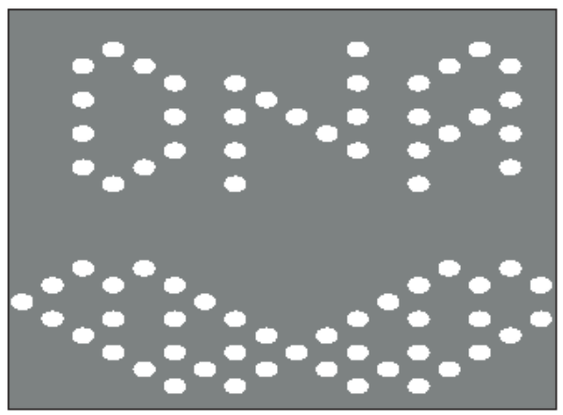


patterns



scaffolded origami



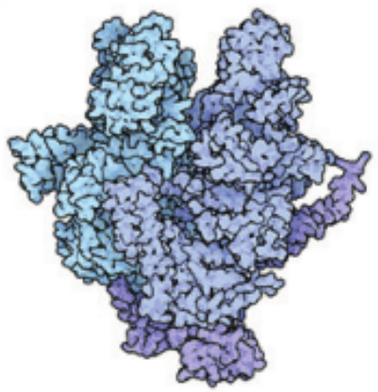


the ribosome

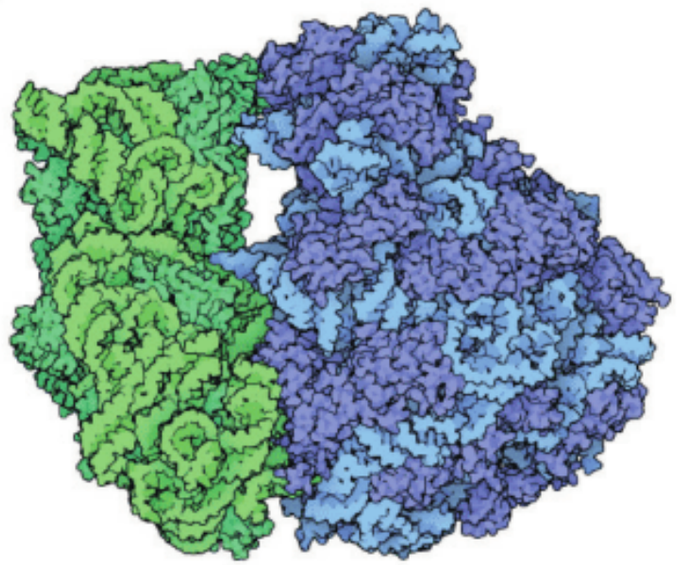
top-down patterning

Going smaller…

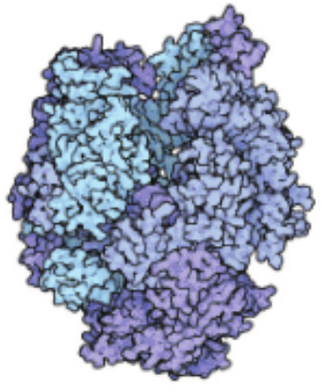
prokaryotic
RNA polymerase (4kmu)



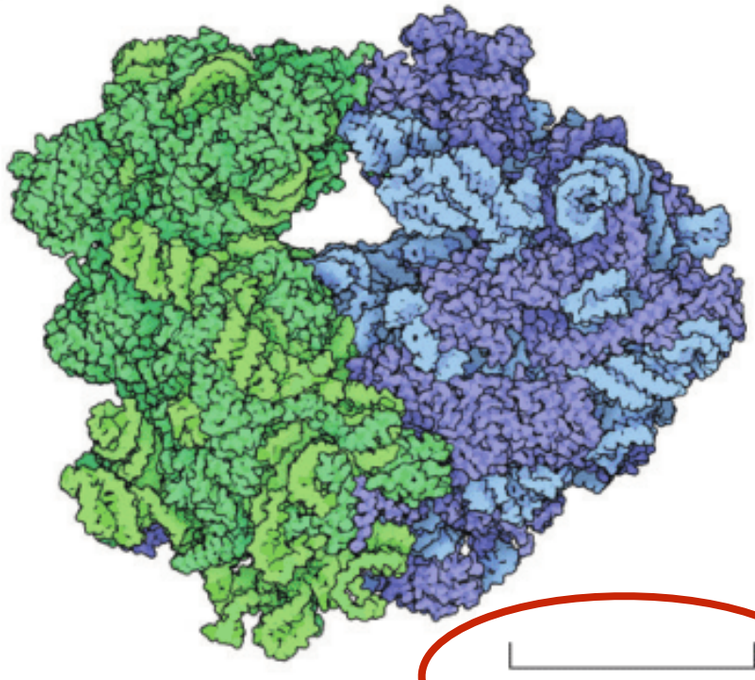
prokaryotic
ribosome (2wdk+2wdl)

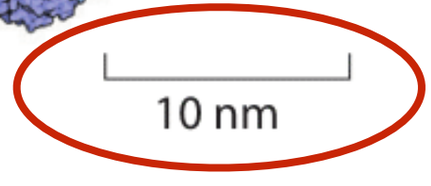


eukaryotic
RNA polymerase (1i6h)

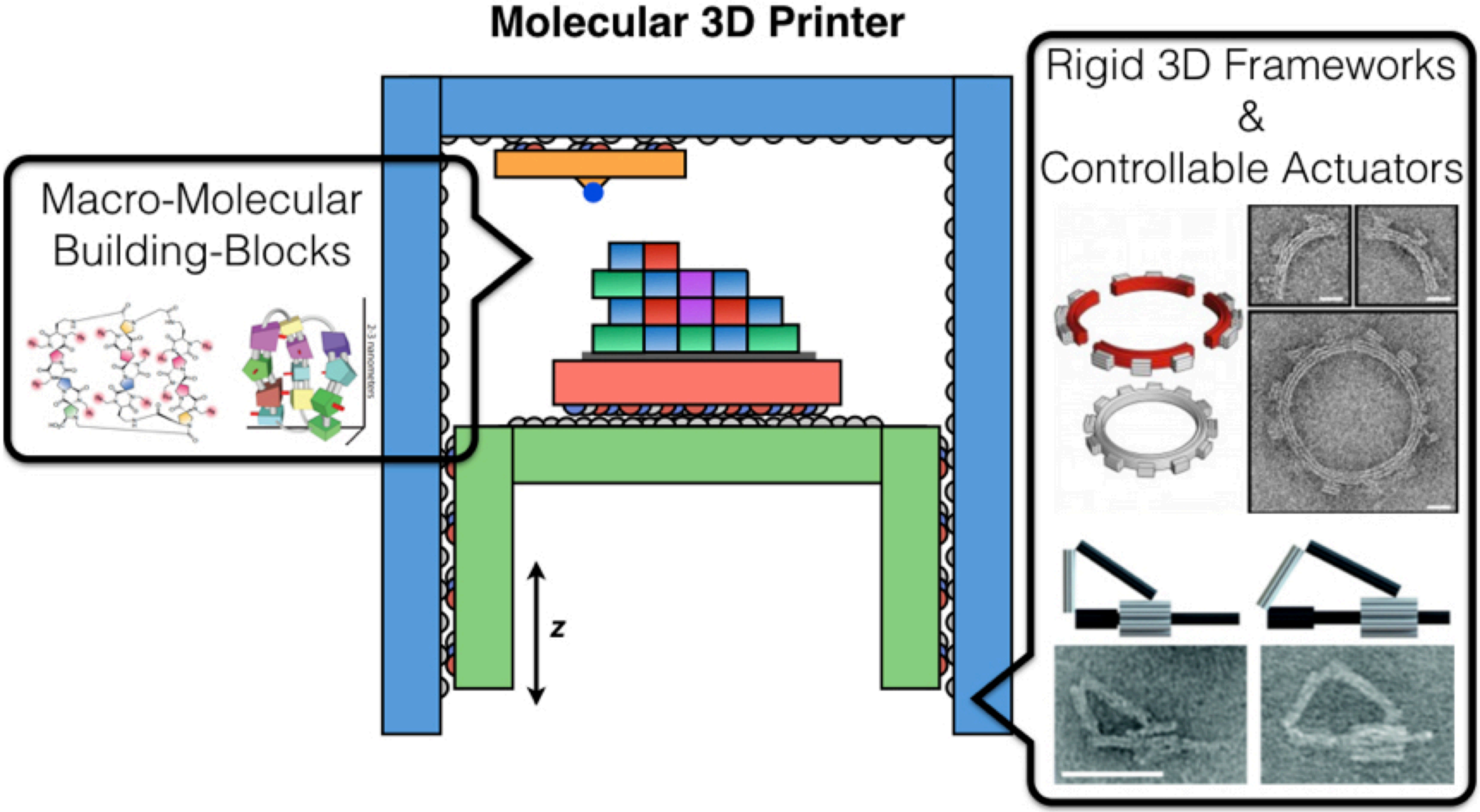


eukaryotic ribosome
(3u5b+3u5c+3u5d+3u5e)



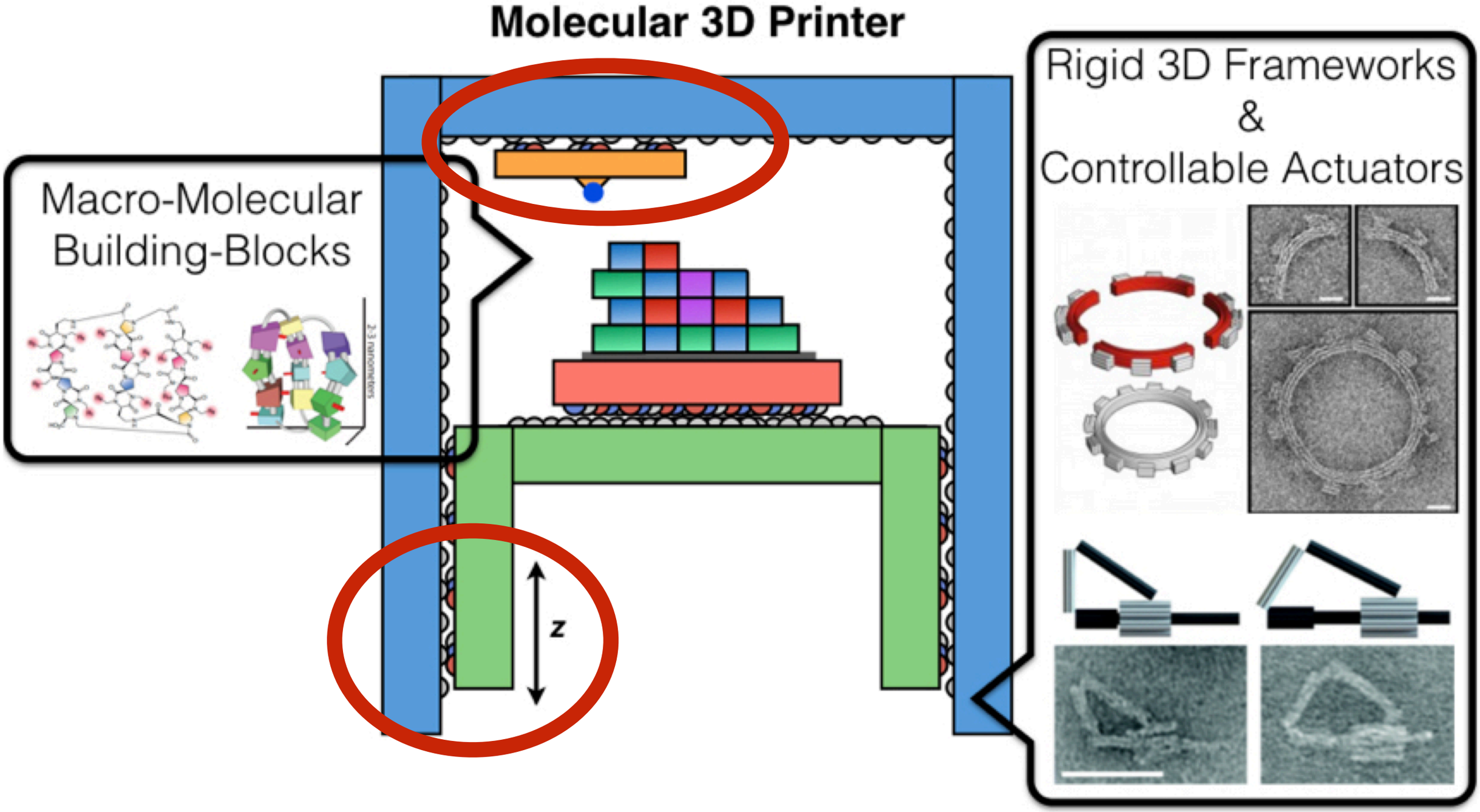


Going smaller…



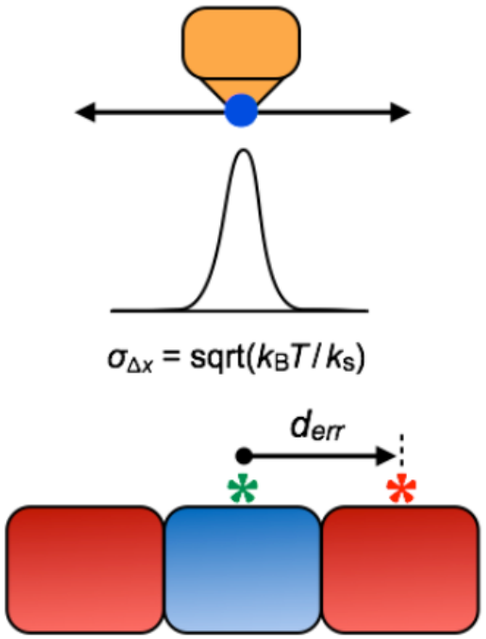
Molecular additive manufacturing workshop, 2016 (w/ Eric)

Going smaller…

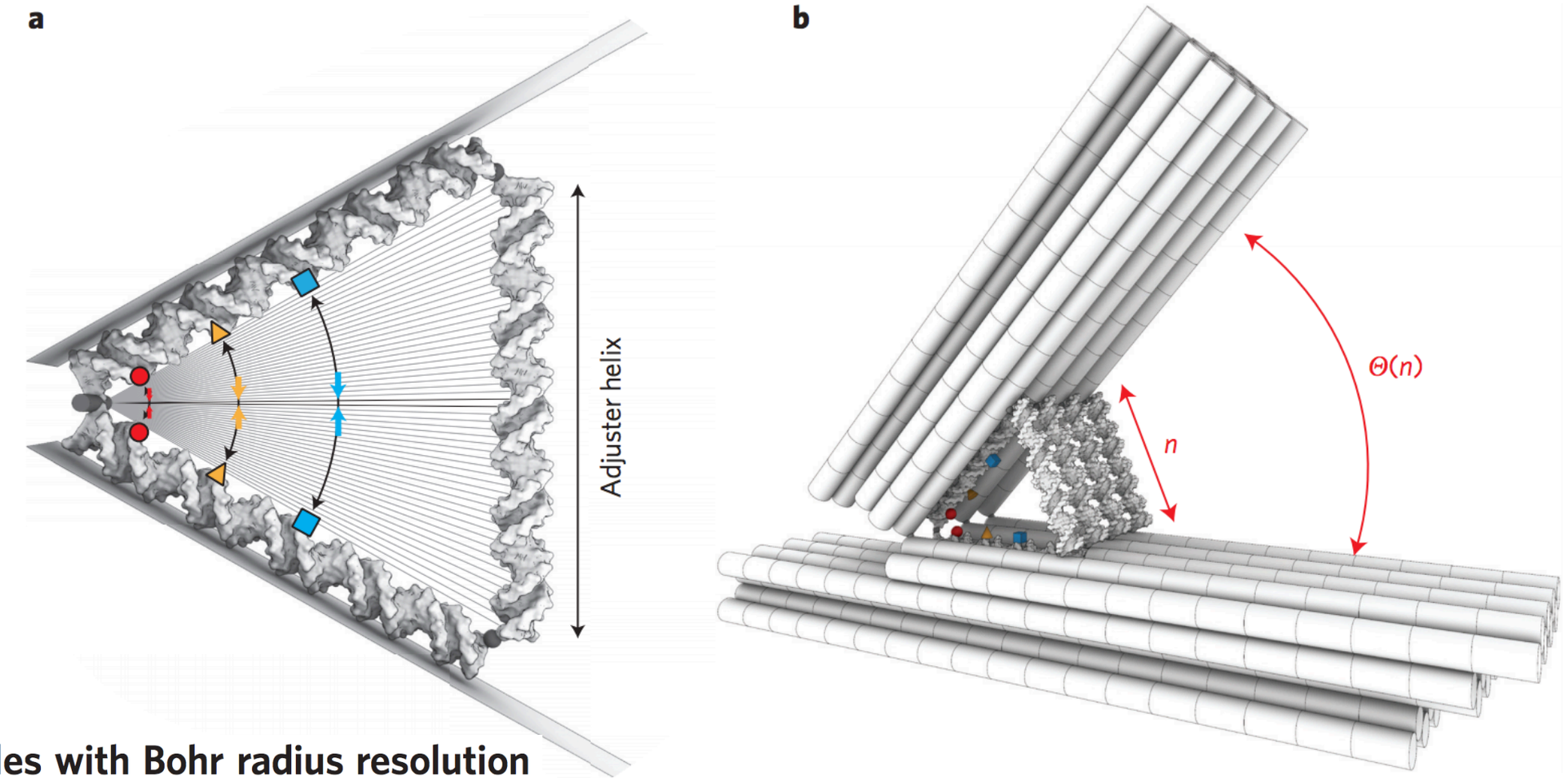


Molecular additive manufacturing workshop, 2016 (w/ Eric)

Going smaller...



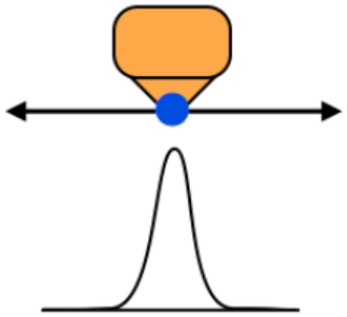
Using this combination of placement and analysis, we rationally adjusted the average distance between fluorescent molecules and reactive groups from 1.5 to 9 nm in 123 discrete displacement steps. The smallest displacement step possible was 0.04 nm, which is slightly less than the Bohr radius. The fluctuation amplitudes in the distance coordinate were also small (±0.5 nm), and within a factor of two to three of the amplitudes found in protein structures[18].



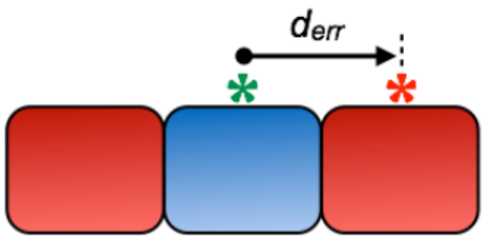
Placing molecules with Bohr radius resolution using DNA origami

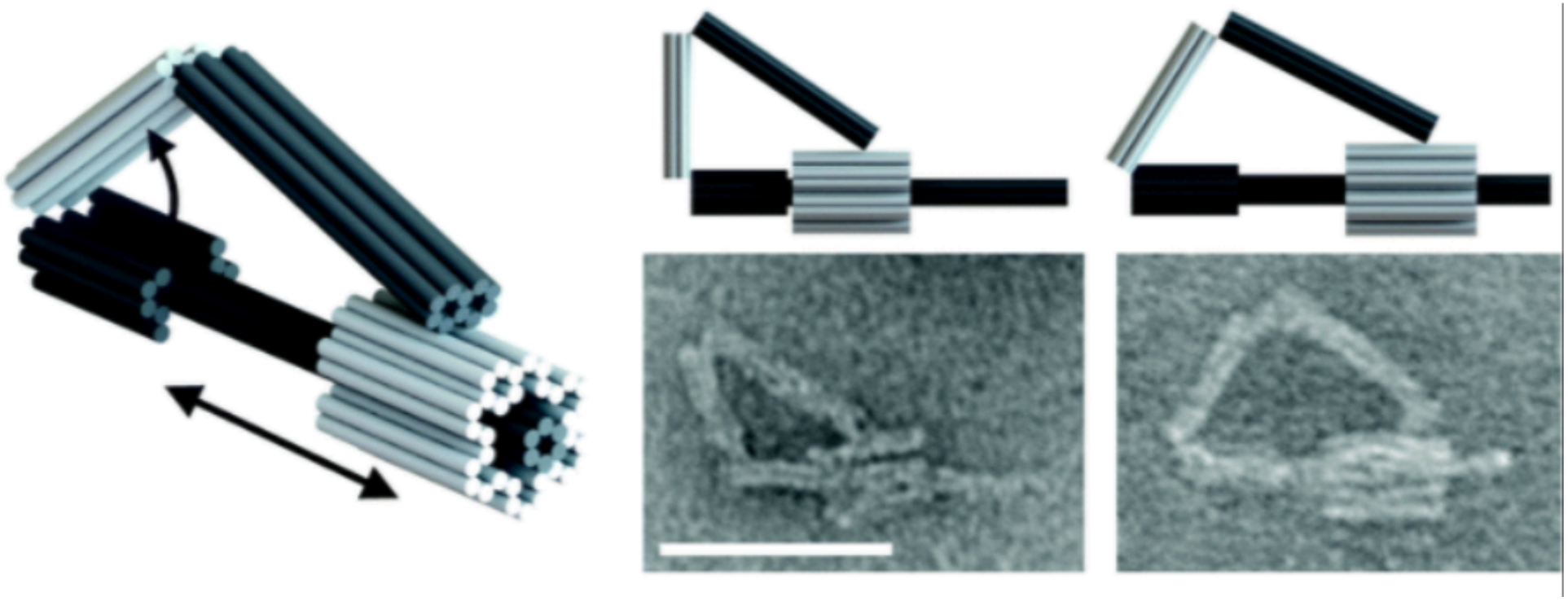
Jonas J. Funke and Hendrik Dietz*

Going smaller…



$\sigma_{\Delta x} = \mathrm{sqrt}(k_B T / k_S)$

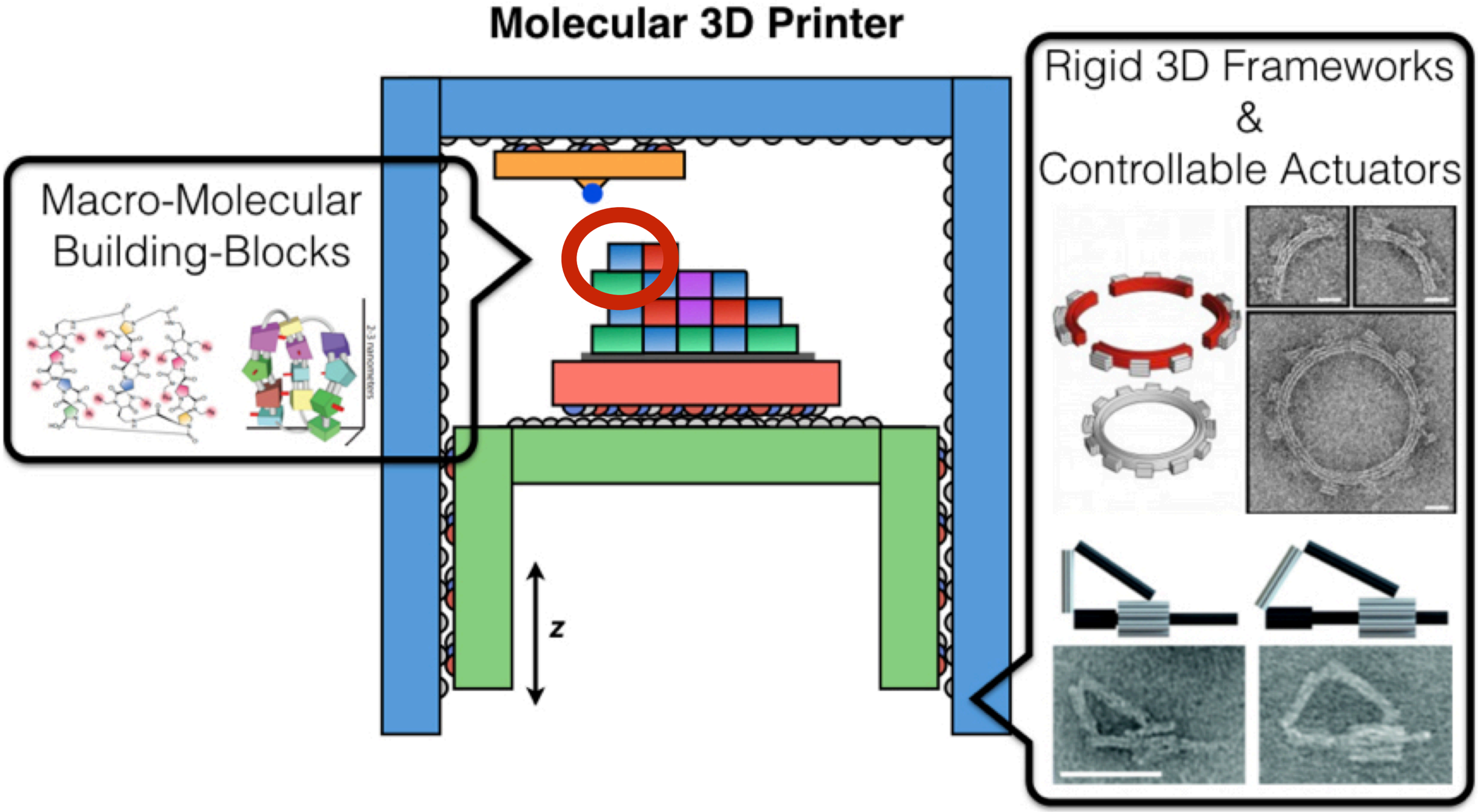




Programmable motion of DNA origami mechanisms

Alexander E. Marras, Lifeng Zhou, Hai-Jun Su, and Carlos E. Castro[1]

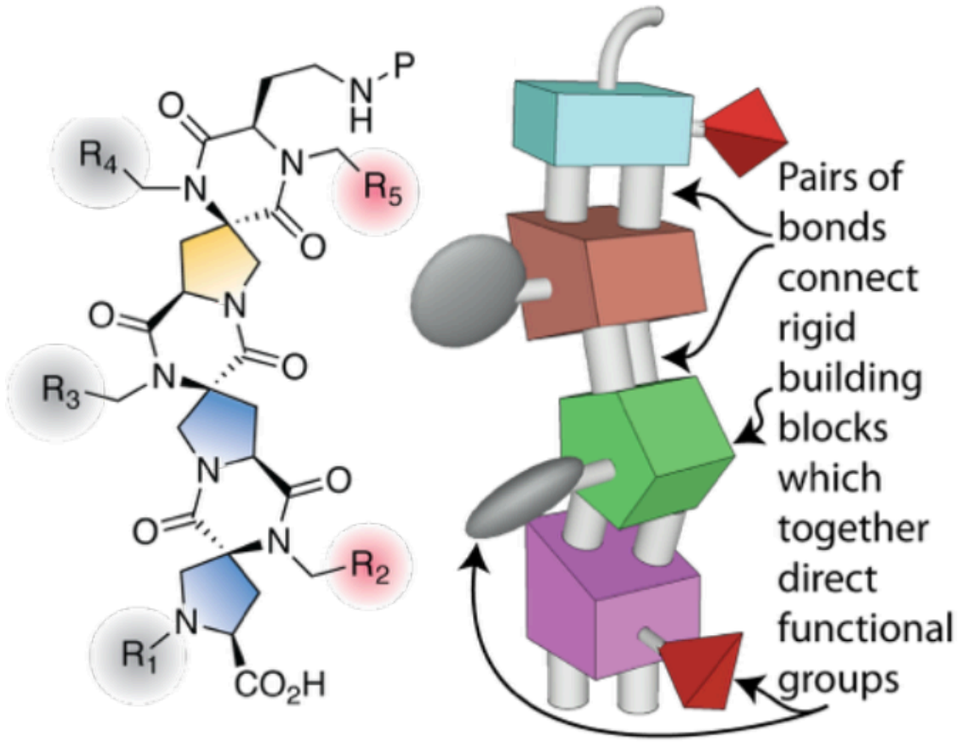
Going smaller…

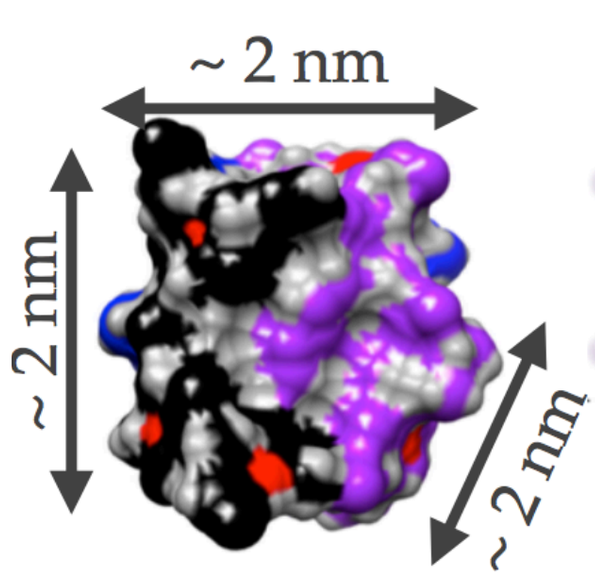


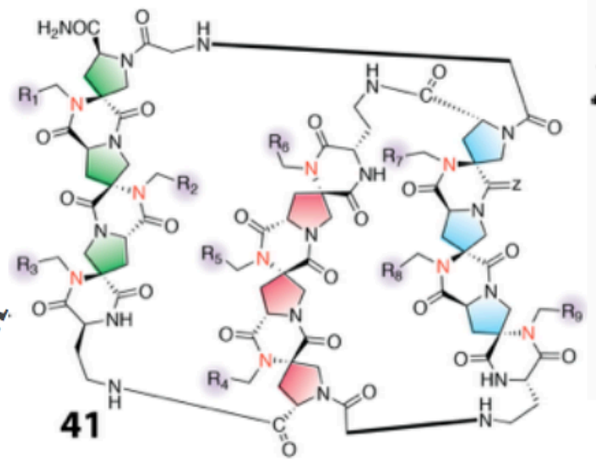
Molecular additive manufacturing workshop, 2016 (w/ Eric)

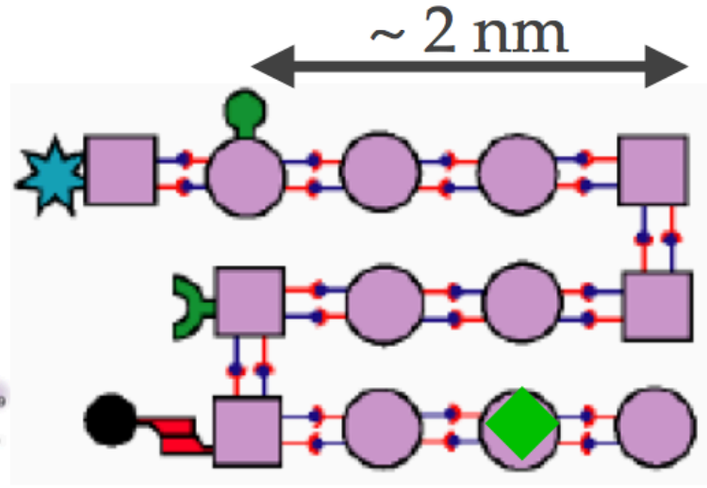
Going smaller…

Spiroligomer "bricks" (Schafmeister):



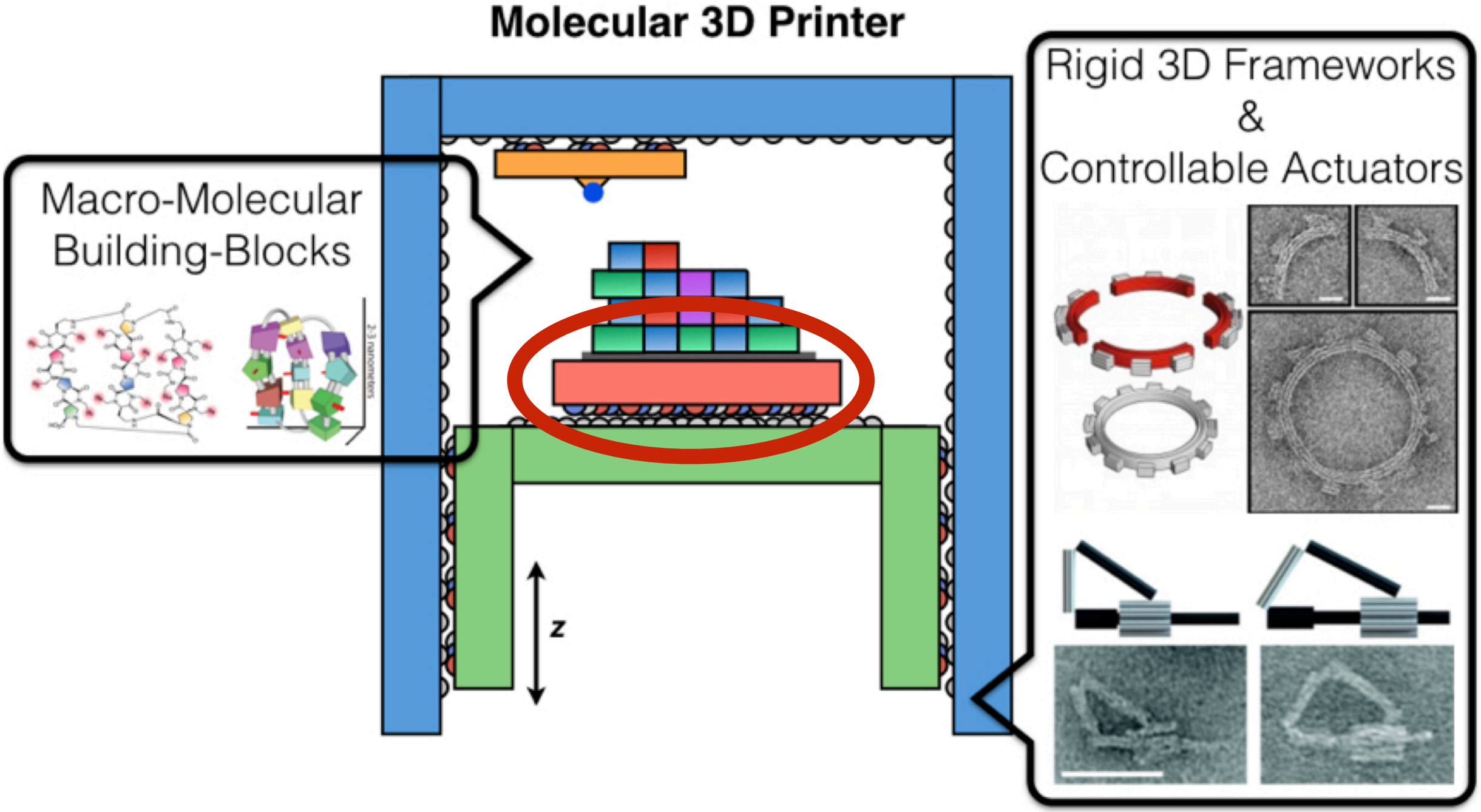






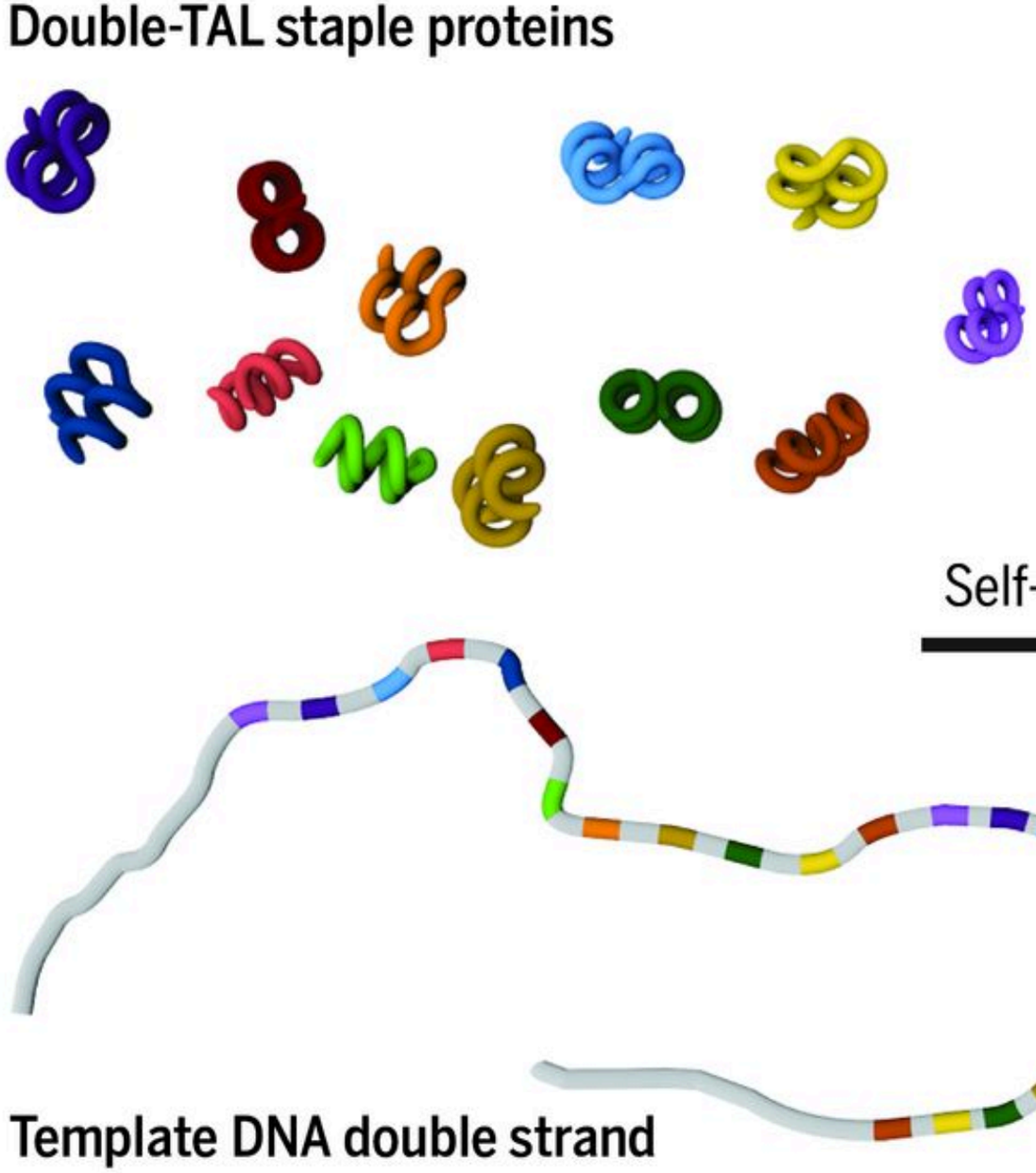
https://www.energy.gov/sites/prod/files/2016/06/f33/Technology%20presentation%20-%20Shafmeister.pdf

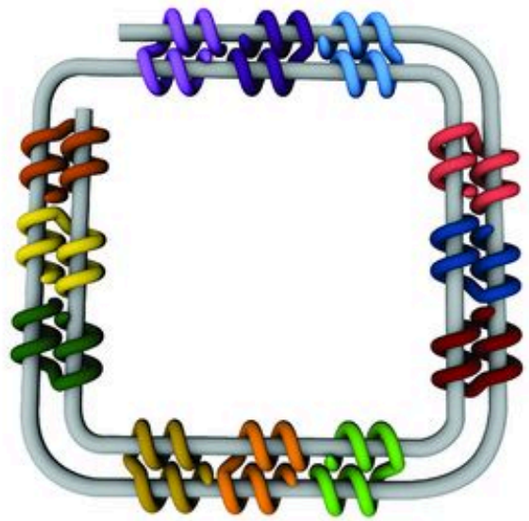
Going smaller…

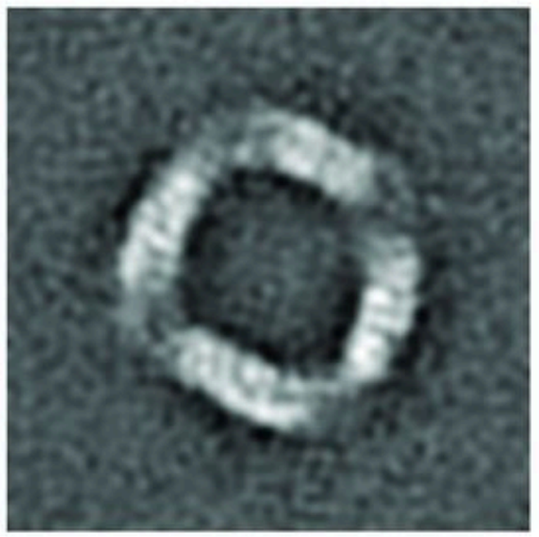


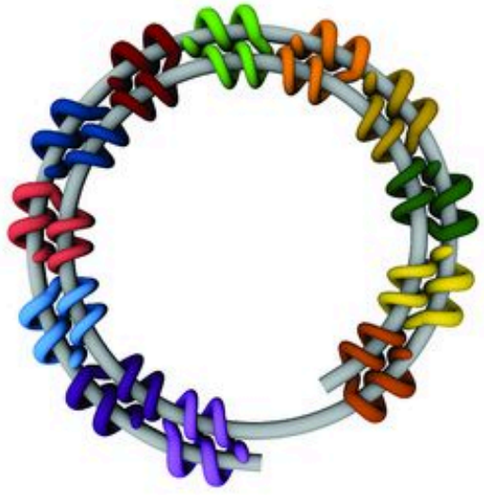
Molecular additive manufacturing workshop, 2016 (w/ Eric)

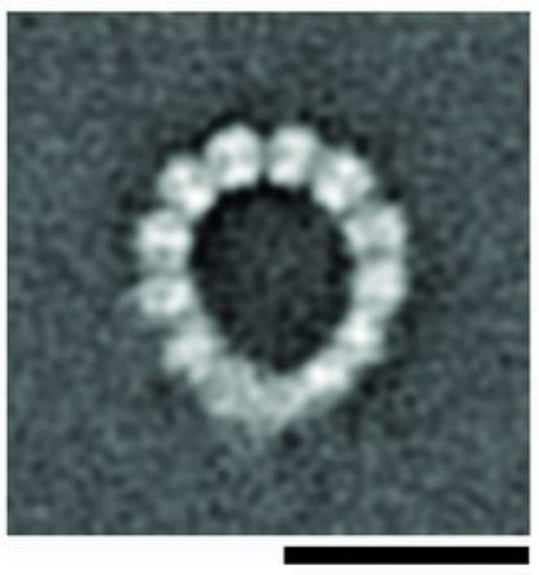
Going smaller…











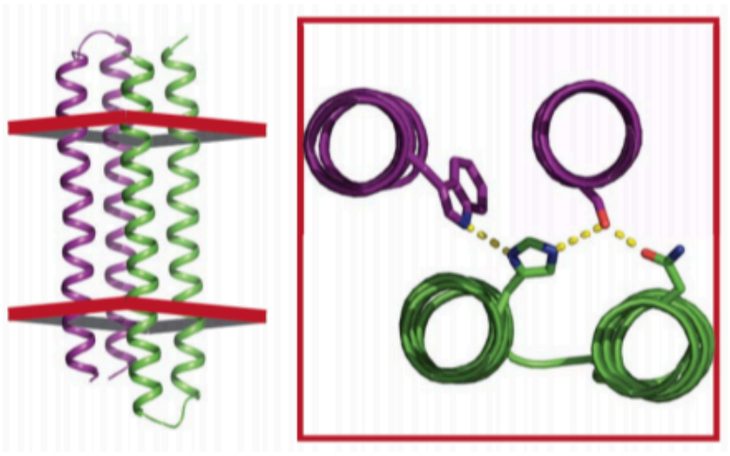
Self-assembly of genetically encoded DNA-protein hybrid nanoscale shapes

Florian Praetorius, Hendrik Dietz*

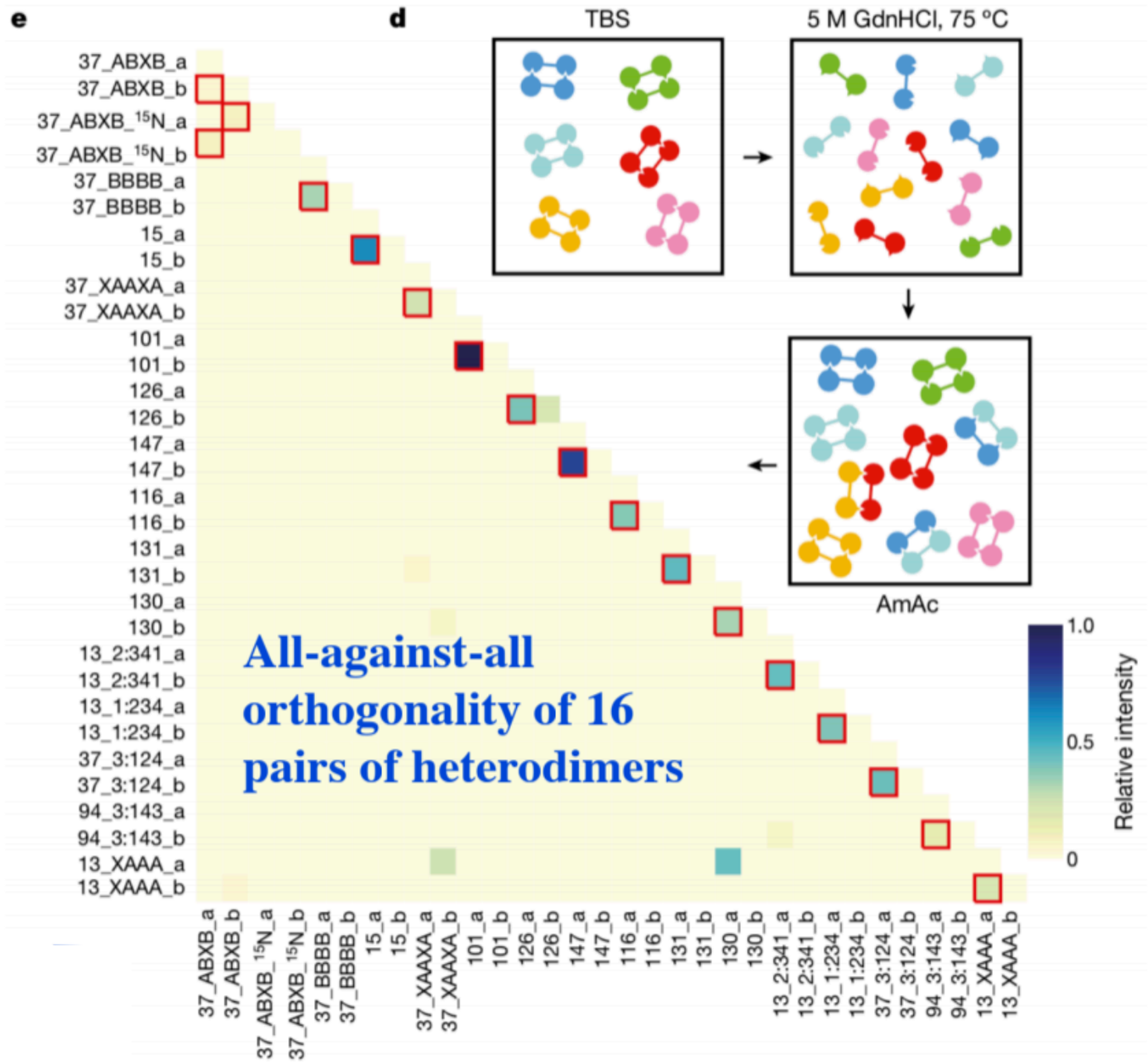
Going smaller…

Design of orthogonal protein heterodimers (Nature, 2018)

…we show that protein–protein **interaction specificity** can be achieved using extensive and **modular side-chain hydrogen-bond networks**.

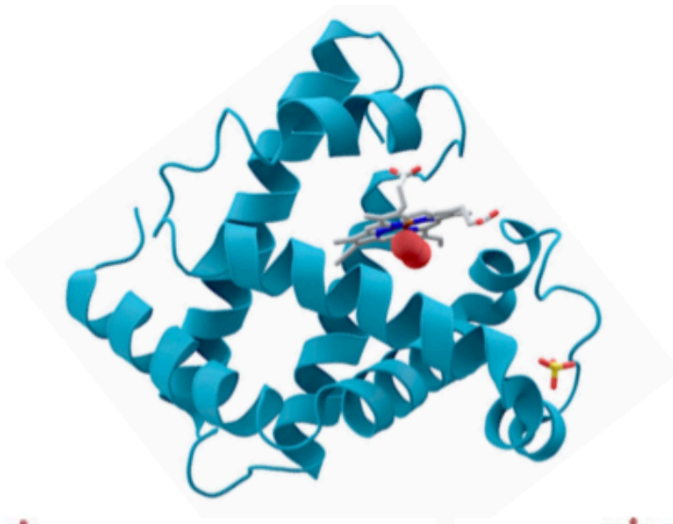


Chen, Z., Boyken, S. E., Jia, M., Busch, F., Flores-Solis, D., Bick, M. J., ... & Bermeo, S. …Baker, D (2019). Programmable design of orthogonal protein heterodimers. *Nature*, *565*(7737), 106-111.

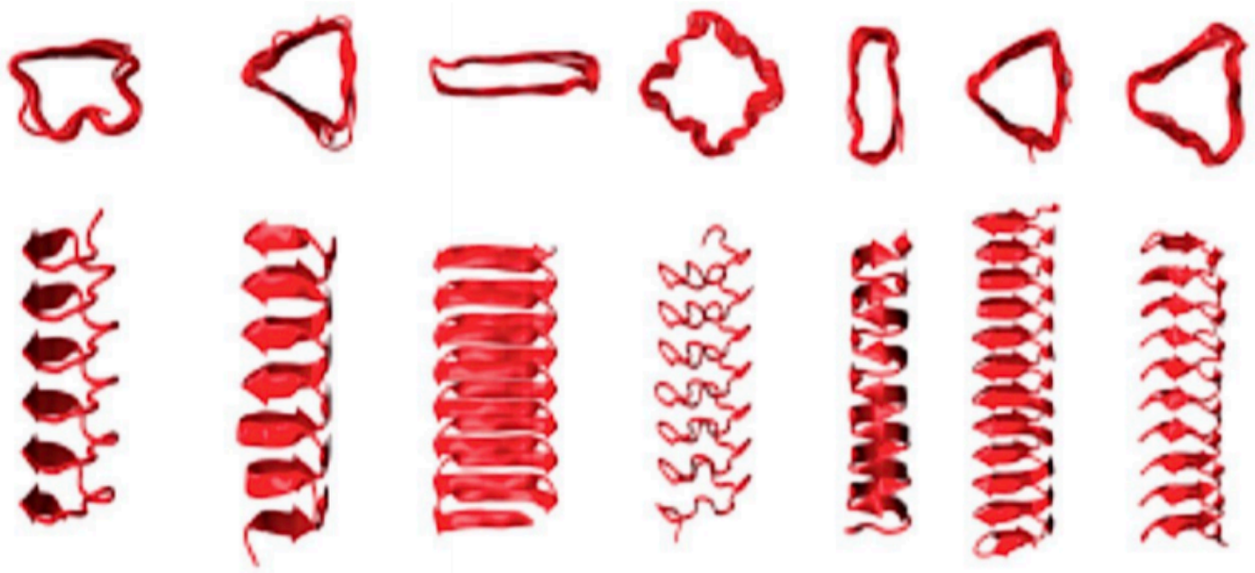


Going smaller... *beta solenoid proteins (BSPs)*

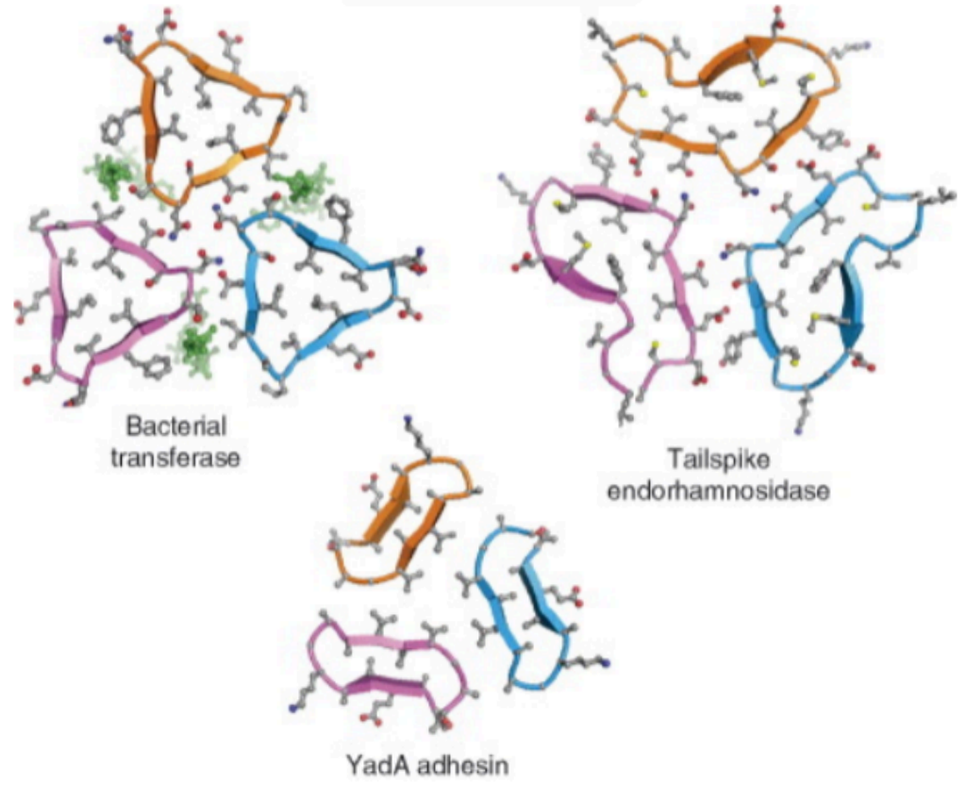
Diverse shapes among natural proteins:



Extensive diversity
even among BSP proteins/domains:

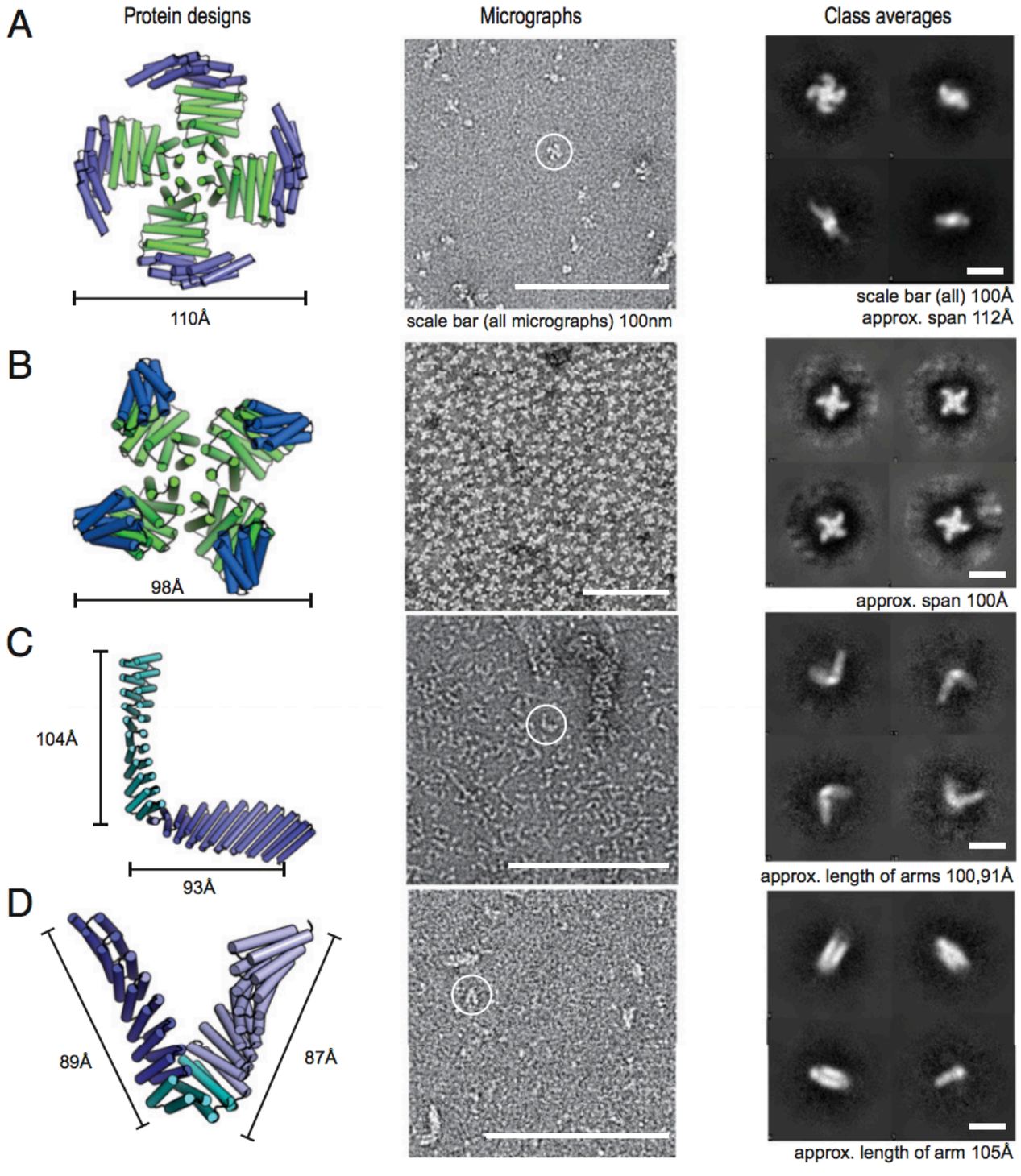


Parker, A., Ravikumar, K., & Cox, D. (2017). Molecular dynamics-based strength estimates of beta solenoid proteins. *Soft matter*, *13*(36), 6218-6226.



Kajava, A. V., & Steven, A. C. (2006). β-rolls, β-helices, and other β-solenoid proteins. *Advances in protein chemistry*, *73*, 55-96.

Going smaller...



Modular repeat protein sculpting using rigid helical junctions

TJ Brunette[a,b,1], Matthew J. Bick[a,b], Jesse M. Hansen[a,c], Cameron M. Chow[a,b], Justin M. Kollman[a], and David Baker[a,b,d]

Summary

Image:

-Using expansion microscopy and in-situ optical sequencing of DNA, we are on path to "*arbitrary resolution, infinite color microscopy*" for bio

-Sequencing neuron-indexing RNA barcodes this way should allow *full mammalian brain connectomes for ~$30M*

Construct:

-DNA origami allows ~5 nm patterning of diverse materials over ~100 nm length scales by self-assembly, with trivial/facile design

-We should be able to organize DNA origami into larger systems by patterning them on *shrunken* DNA microarrays

-A "*molecular 3D printer*", made through various steps involving DNA origami, DNA templating of proteins / peptides, and modular protein engineering with fixed backbones and *programmable interfaces*, could be an interesting first step to explore and demo "APM" principles