OUR NEED FOR NANOFABRICATION

BOB AUSTIN
JONAS TEGENFELDT
HAN CHAO
RICHARD HUANG
STEVE CHOU
JIM STURM
TED COX
HAROLD CRAIGHEAD (CORNELL)
MARIO CABODI (CORNELL)
TOM DUKE (CAMBRIDGE)

THANKS TO PAL ORMOS FOR INVITING SHIRLEY AND I.

I WILL TALK ABOUT TECHNOLOGY THAT CAME FROM COMPUTERS. BUDAPEST, HUNGARY IS THE BIRTHPLACE OF JANCSI VON NEUMANN, THE FATHER OF THE STORED CODE COMPUTER. THIS IS HOLY GROUND!

ACTUALLY, PRINCETON HAS HAD A LONG EXPERIENCE WITH HUNGARIANS ("Martians"): von Neumann, Wigner and many on our staff. A bank near us is Magyar Savings. Hungarians have a long and glorious association with biological physics:

http://www.kfki.hu/fszemle/archivum/fsz9905/vicsek.html

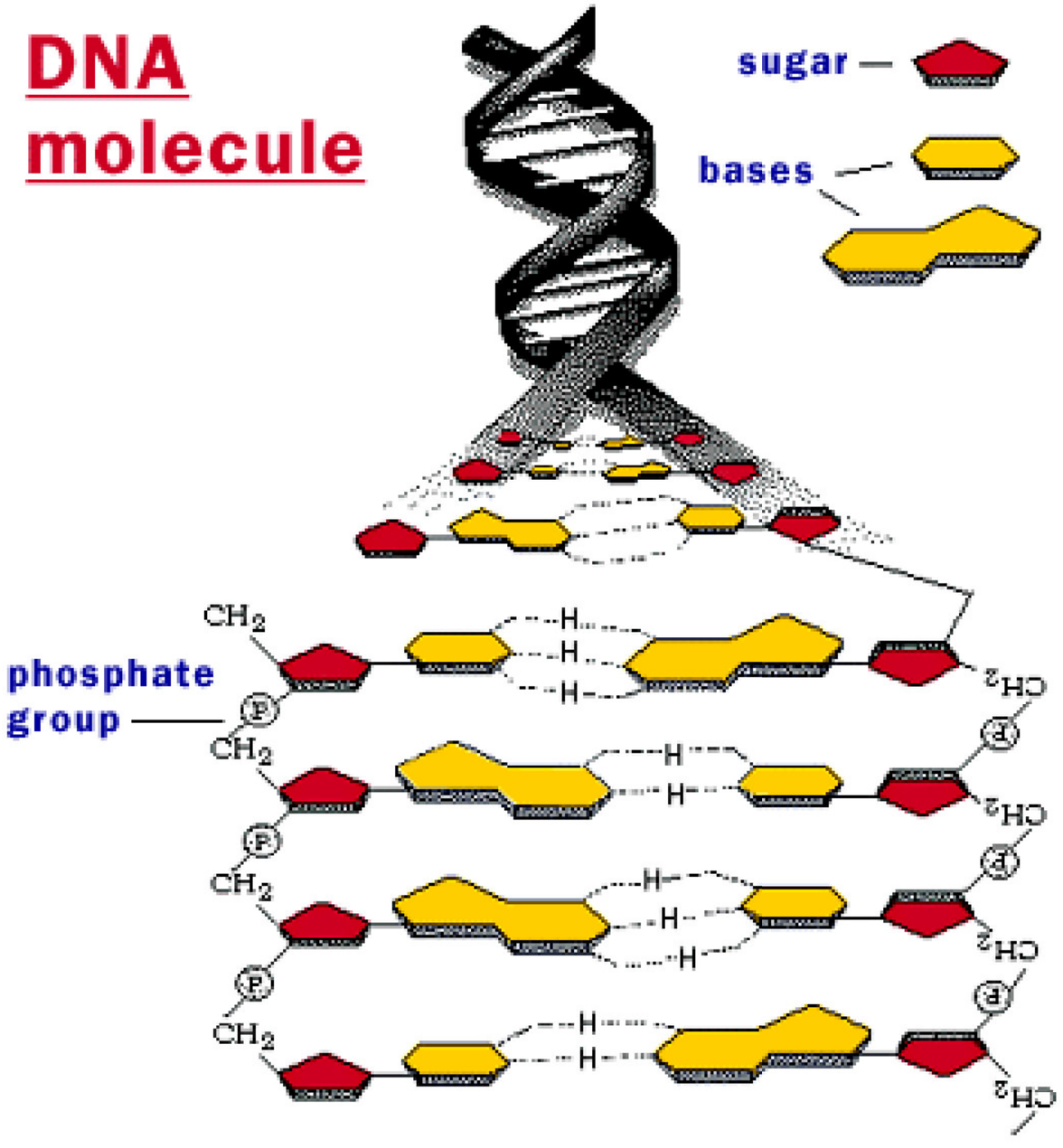
I WANT TO TALK A LITTLE BIT ABOUT HOW THE TECHNOLOGY DEVELOPED BY THE COMPUTER INDUSTRY CAN (WILL) CHANGE THE FACE OF MODERN BIOLOGY.

THIS IS JUST A SMALL SNIPPET OF THINGS BEING DONE IN MY LAB, FROM BACTERIA CHEMOTAXIS TO PROTEIN DYNAMICS.

The integrated circuit revolution, made possible by microfabrication technology, is just now entering the world of biology. I am old enough to have built a Heath Kit tube tuner and amplifier, and to notice that the transfer characteristics of a FET resemble a triode. So, I know the revolution that can happen.

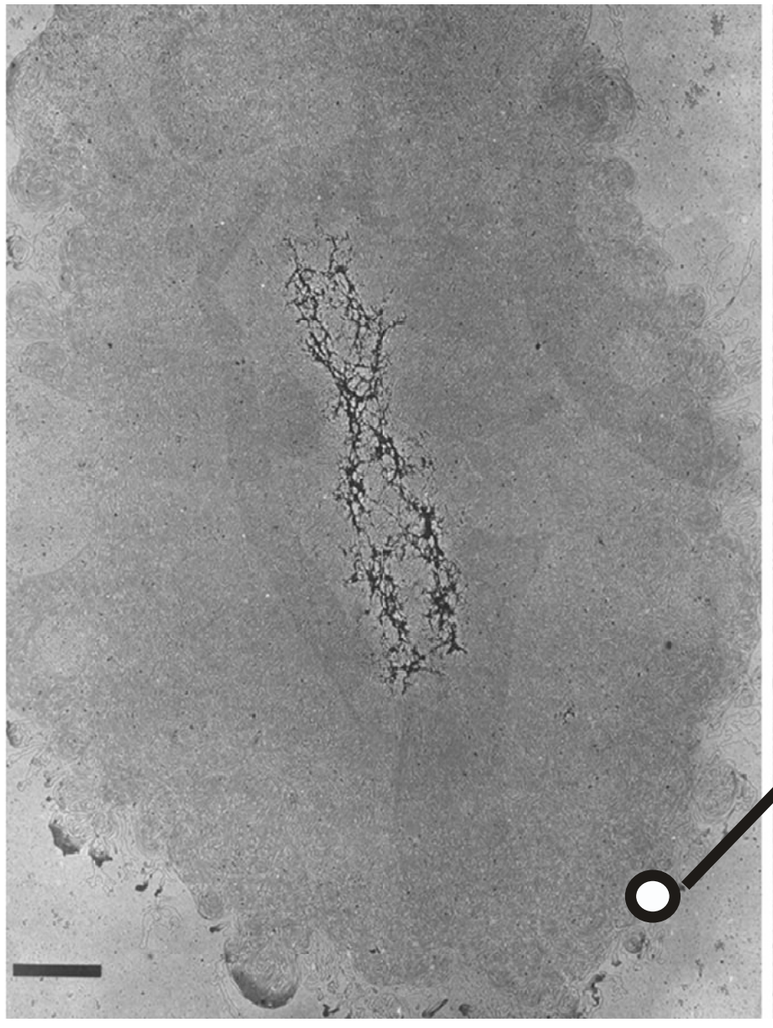
The world of biology is inherently on the micron and below scale, and this is where microfabrication lives. We can process, examine, move biological objects at their natural length scale. Further, the world of biology is all about heterogenity: no two objects are alike. Sometimes, the rarest one is the most interesting. Microfabricated devices, which basically are "flatlanders" can find those rare ones.

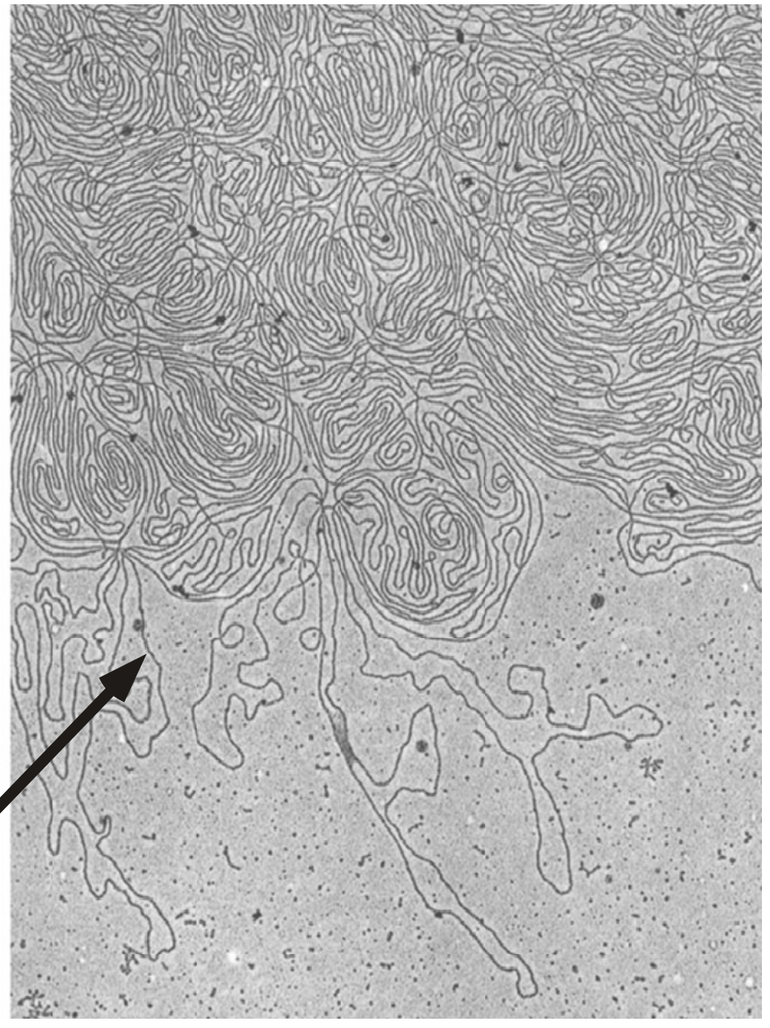
DNA molecule



DNA contains the code for you. There are about 3 billion basepairs in the human genome, or about 1 meter of DNA in each cell.

In total, enough DNA to reach from the sun to Pluto.





Sequencing the human genome was an absolutely stupendous problem.

The Human Genome Project and Celera Genomics Corporation have both completed an initial sequencing of the human genome - the genetic blueprint for human beings. Clinton congratulated the scientists working in both the public and private sectors on this landmark achievement, which promises to lead to a new era of molecular medicine, an era that will bring new ways to prevent, diagnose, treat and cure disease.

This enormous project was done about 3x faster then the most optimistic estimates in the early 90's and for about 1/10 the estimated price. Basically brute force roboticsdid it and computer developments, some knowledge of physics

Stanley Fields article in Science:

"If the architect you hired to design your home brought you a blueprint that solely consisted of a long list of parts that began "windowwabeborogovestaircasedoorjubjub...," you might start to wonder if and when you will see your new house. Some people have similar reservations about the recently "completed" human genome sequence, heralded as the "genetic blueprint" that will revolutionize biology and medicine. Deciphering how a mere 10^7 nucleotides result in a yeast cell--let alone how 3×10^9 nucleotides result in Tiger Woods or Britney Spears--cannot begin until the genes have been annotated. This step includes figuring out the proteins that these genes encode and what they do for a living. But understanding how all of these proteins collaborate to carry out cellular processes is the real enterprise at hand.

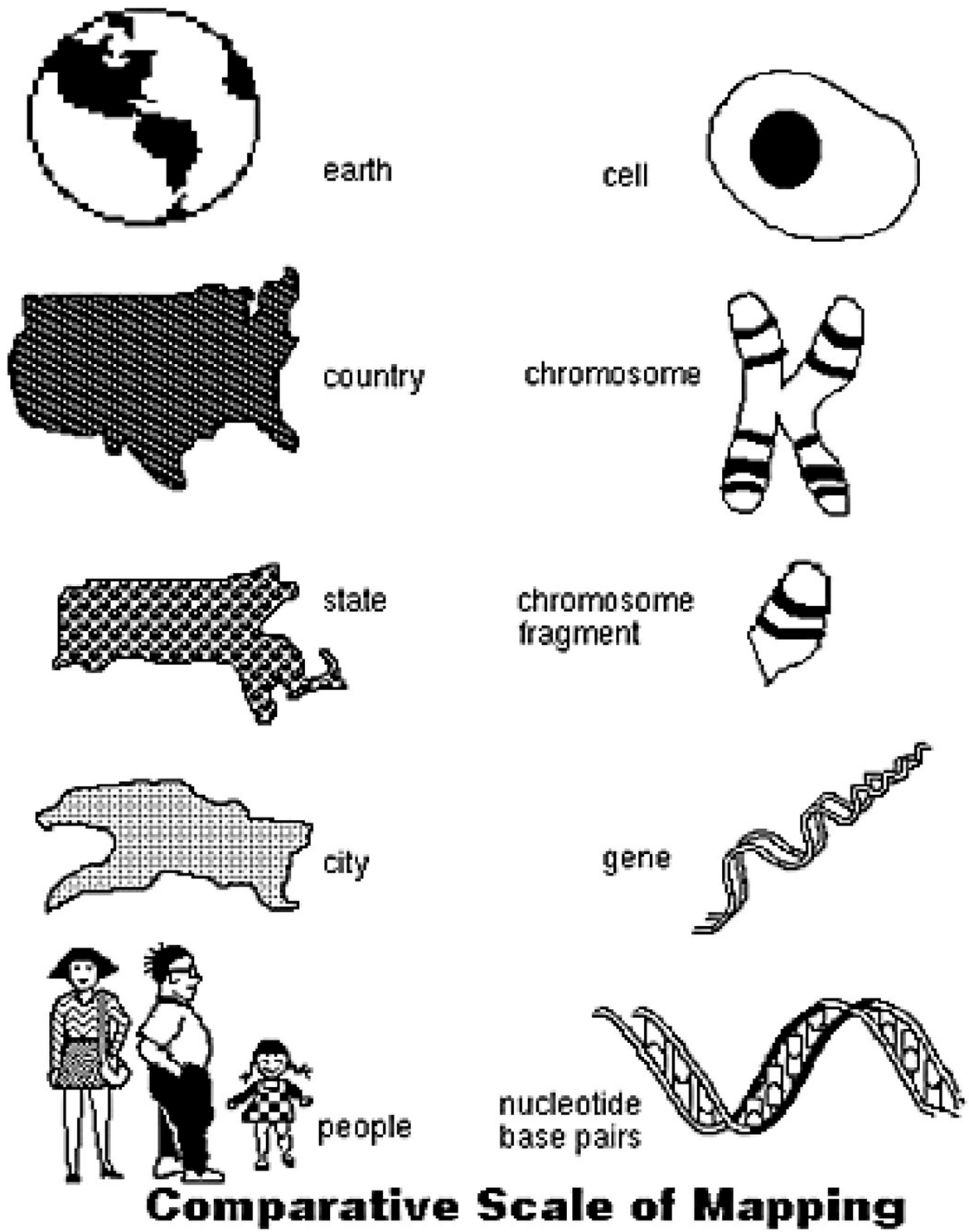
"SEQUENCING AND RESEQUENCING THE HUMAN GENOME": A RECENT MEETING AT THE NIH.

THERE WAS A CLEAR CONCENSUS: WE NEED TO DROP THE PRICE OF SEQUENCING A LARGE GENOME FROM ABOUT $100,000,000 TO $1,000: A DROP IN 5 ORDERS OF MAGNITUDE.

HOW CAN WE DO IT?

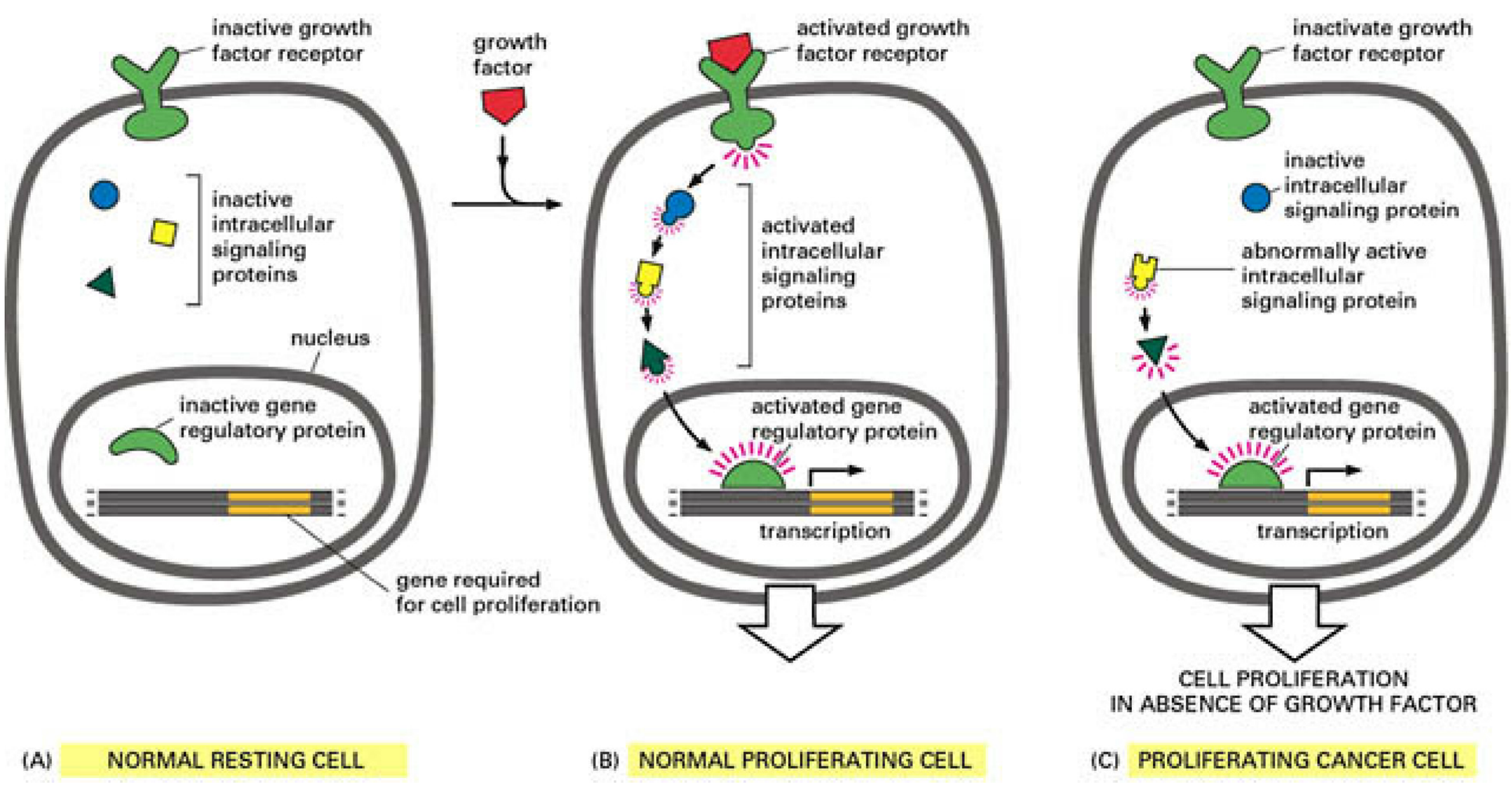
EVEN AS WE WORK ON LEARNING HOW TO DO THE $1,000 GENOME, THERE ARE ADDITIONAL QUESTIONS BEYOND GENOMICS, FOR GENOMICS BASICALLY ONLY TELLS ONE THE BLUEPRINT OF A BUILDING, NOT THE ACTUAL FUNCTIONING STATE OF THE COMPLEX. THE ANALOGY TO A BUILDING IS NOT A GOOD ONE, FOR BIOLOGICAL SYSTEMS ARE FAR MORE COMPLEX.

EPIGENETICS

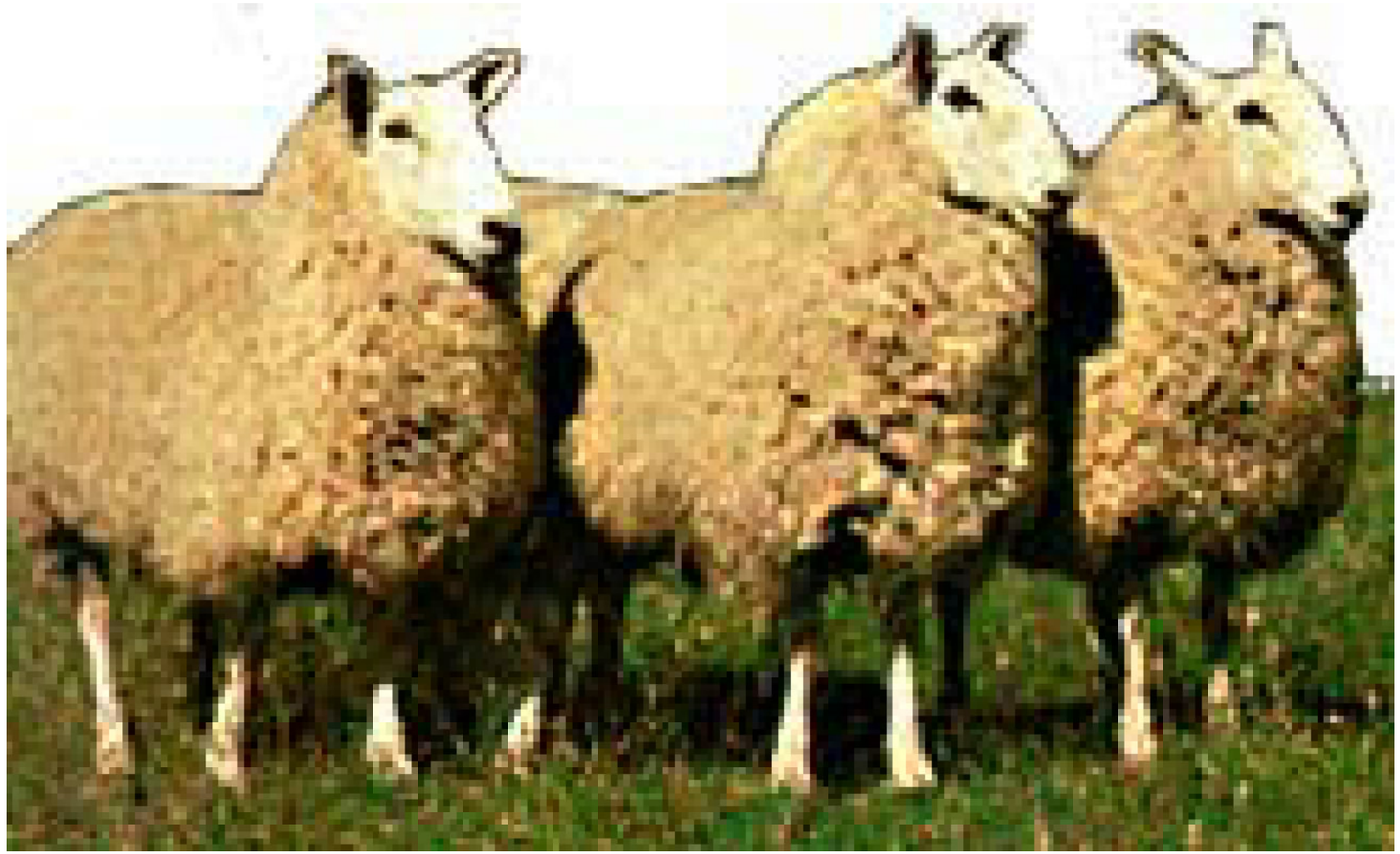


Comparative Scale of Mapping

Nearly all the 10^{14} cells in your body contain the SAME genome. Probably each cell has a slightly different genetic sequence. But, it isn't the genome that makes each cell different.



Promoter and repressor proteins, which bind to specific parts of the genome, control expression. The CYTOPLASM of the cells contains the DYNAMIC control information. The DNA is the ROM, the proteins are the OS that makes a liver cell a liver cell.



That's why there was all the excitement about Dolly: they took the NUCLEUS from a fully differentiated cell in the udder, put it into the egg cell of another sheep which presumably had the right protein content to reset the clock, and transformed a mature cell into a "fertilized" egg.

Cells are dynamic entities, they are constantly changing with time as the epigenetic state of the cell changes, that is, as the pattern of expressed proteins changes.

Chip-based hybridization technologies CAN map out this critical time dependence, but only with great effort and expense since the technology is very indirect.

A more dramatic and sinister aspect of the changes in a cell are the transformations that occur with accumulated genetic mistakes. Critical "gateway genes" that control fundamental metabolic and control pathways become mutated and the cell becomes unstable and starts unregulated growth.

This is a highly complex and coordinated process.

$L_{n+1} = G\,(1 - L_n)\,L_n$ (the logistic equation)

This equation shows chaotic behavior as $G>3.3$
This can be an epigenetic chaos.

In the case of metastatic cancer, the transformed cells that have lost control and think they are immortal spread through the blood stream and form little colonies. These transformed cells are different than normal cells:

(1) genome is different (mistakes)
(2) control proteins are different (wrong controls)
(3) cytoskeleton is different (growing too fast)
(4) surface proteins different (not a differentiated cell)

The problem is, it only takes 1 transformed cell to kill you. You need to FIND and CHARACTERIZE those rare cells.

MY LAB AND COLLABORATORS ARE DOING A WIDE RANGE OF PROJECTS. LET ME (INEPTLY) DISCUSS TWO PROJECTS THAT INVOLVE NANOLITHOGRAPHY:

(1) FRACTIONATING DNA ON A CHIP

(2) SCANNING THE PROTEIN COMPONENTS BOUND TO A SINGLE DNA MOELCULE

I'LL TELL YOU RIGHT NOW THAT I AM DEPRESSED HOW SLOWLY THE WORK GOES, I LACK THE CHARISMA AND SCIENCE SKILLS TO PUSH THIS WORK FORWARD AT WARP SPEED

SEQUENCING DNA USING SEMI-CONVENTIONAL TECHNOLOGIES REQUIRES BEING ABLE SEPARATE DNA MOLECULES THAT ARE 0.3 NM (1 BASEPAIR) DIFFERENT IN LENGTH.

PRESENT TECHNIQUES CAN DISTINGUISH BASEPAIRS OUT TO ABOUT 700 BP. PRESENT MACHINES CAN DO THIS AT ABOUT 1 BASEPAIR/SEC.

IT'S GREAT, BUT NOT GOOD ENOUGH. CAN WE DO BETTER USING NANOFABRICATED TECHNIQUES?

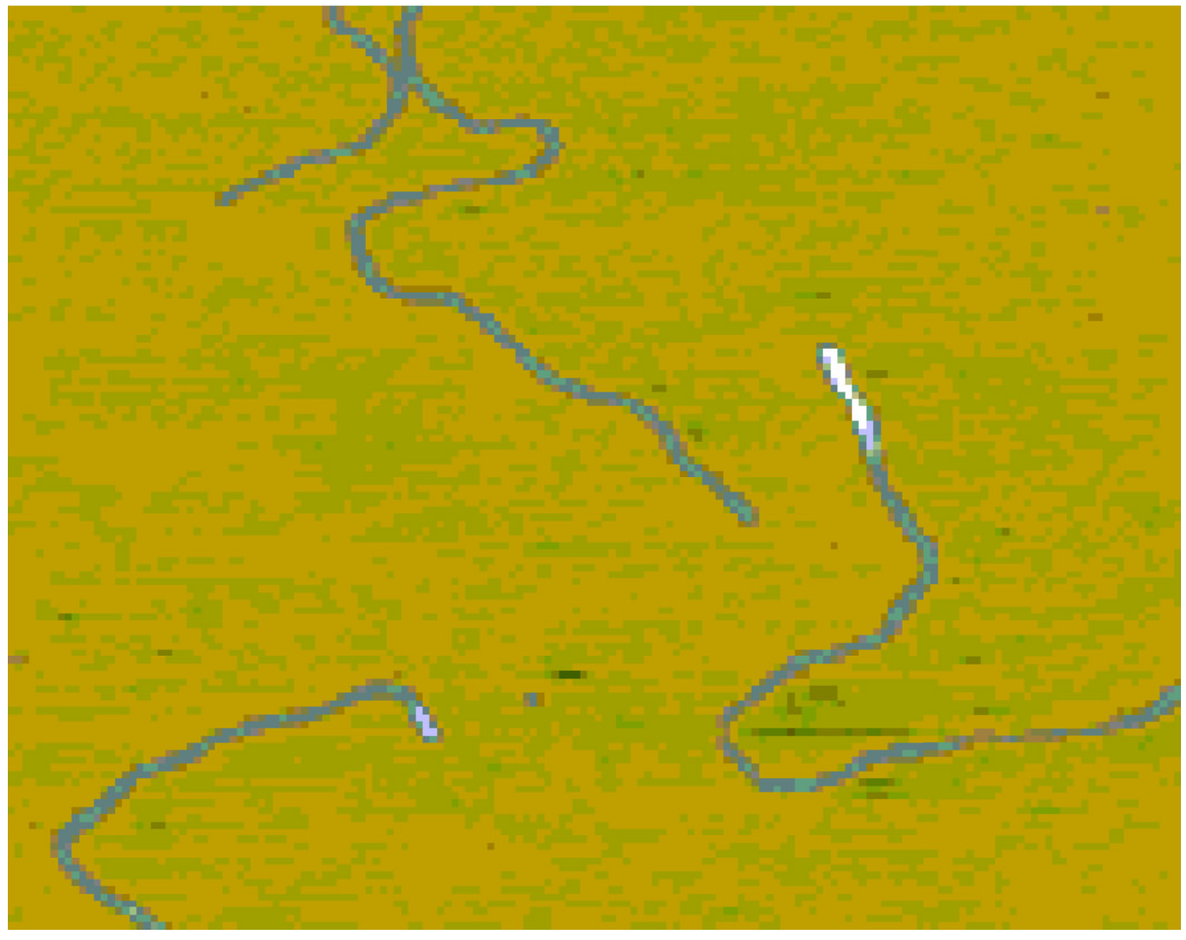
DNA:

(1) A VERY long polymer, up to cm long!

(2) About 15 Angstroms in diameter

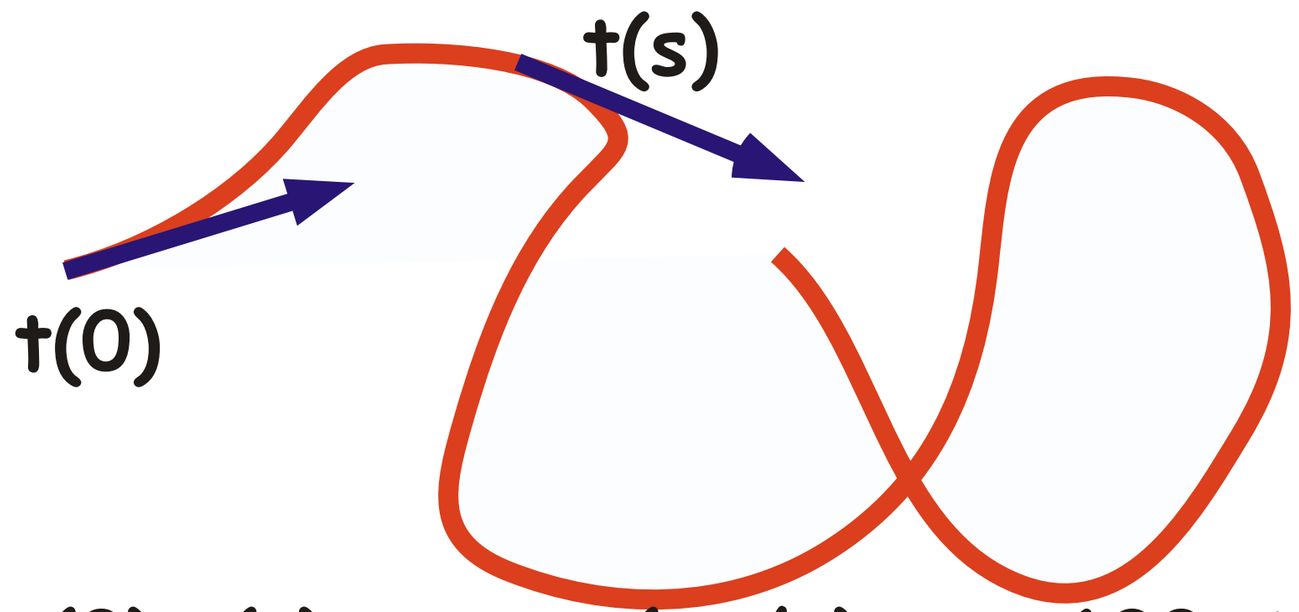
(3) Negatively charged! 2e- per basepair

(4) "Flexible"



AFM pictures of DNA fragments, Paul and Helen Hansma

A very important concept here: the persistence length "p" of a flexible polymer. Basically, it is a measure of how far you move along an arc before thermal energy bends the polymer randomly.



<t(0) t(s)> = exp(-s/p) p = 600 Angstroms = 200 bp (double-strand DNA). single-stranded DNA far smaller p: about 5 nm

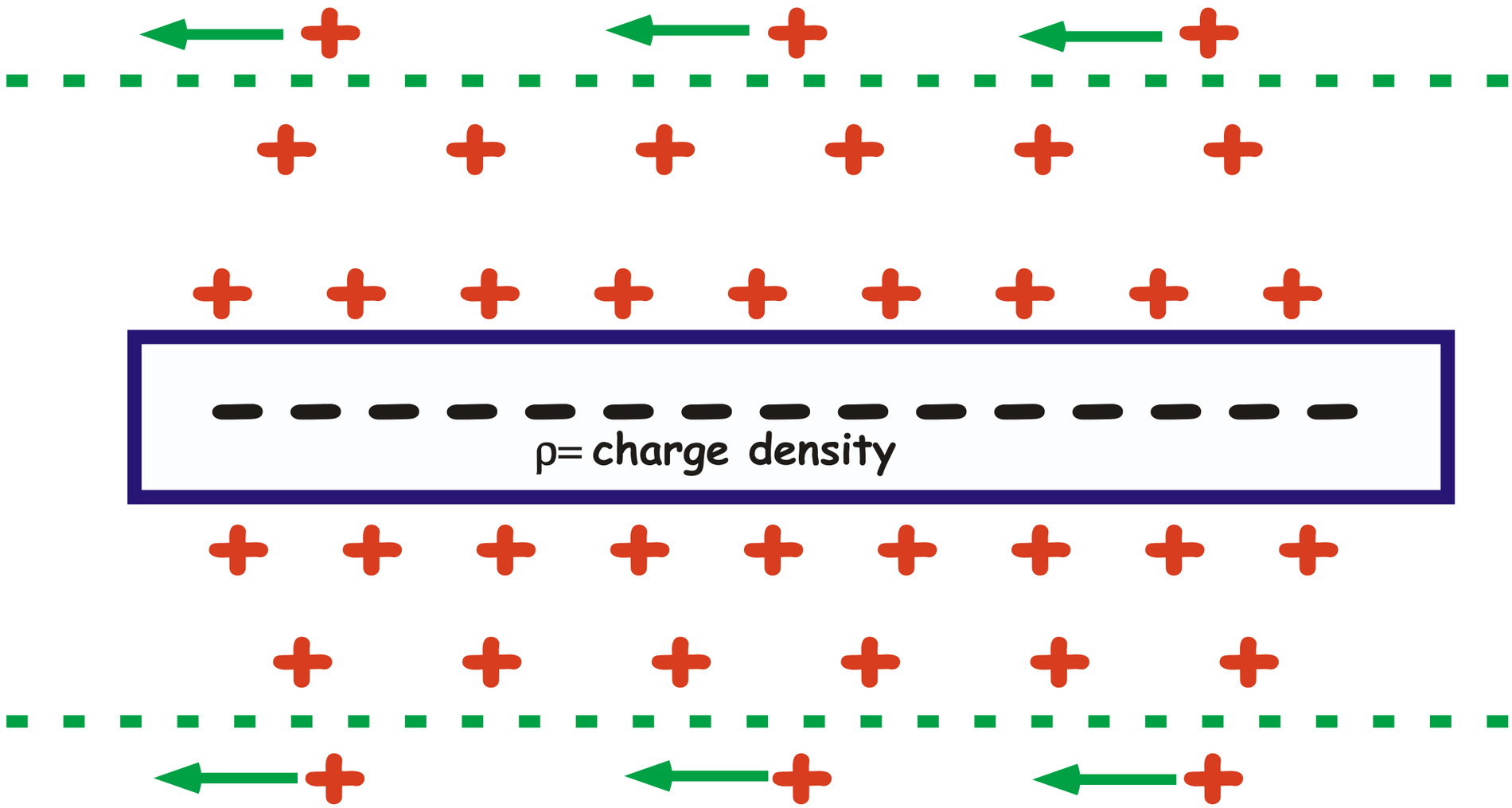
This number controls statistics, dynamics

We transport objects in our chips using primarily two forces, hydrodynamic and electrophoretic. I want to stress right away that the two forces are very closely connected because of the counter-ion shielding of polyelectrolytes in solution. This means that ions moving in the solution pull via hydrodynamic drag on the polyelectrolyte and this shearing action basically cuts off long-range hydrodynamic interactions. This has profound consequences.

This really old video, from my old student Wayne Volkmuth, explains as best as I can how you can go very wrong without realizing the connection between electrophoresis and hydrodynamics.

an old, old movie showing the limitations of the american mind

Electrophoresis: really hydrodynamics because water is an isulator. The moving ions pull molecules along via hydrodynamic coupling.



Charged objects are neutral in a salt solution in the absence of a flow of ions. The moving ions "shear" away bound counter ions that are held at a potential energy less than some characteristic value known as the Zeta potential.

The result is that the polymer is "free draining" and the parts are decoupled from each other.

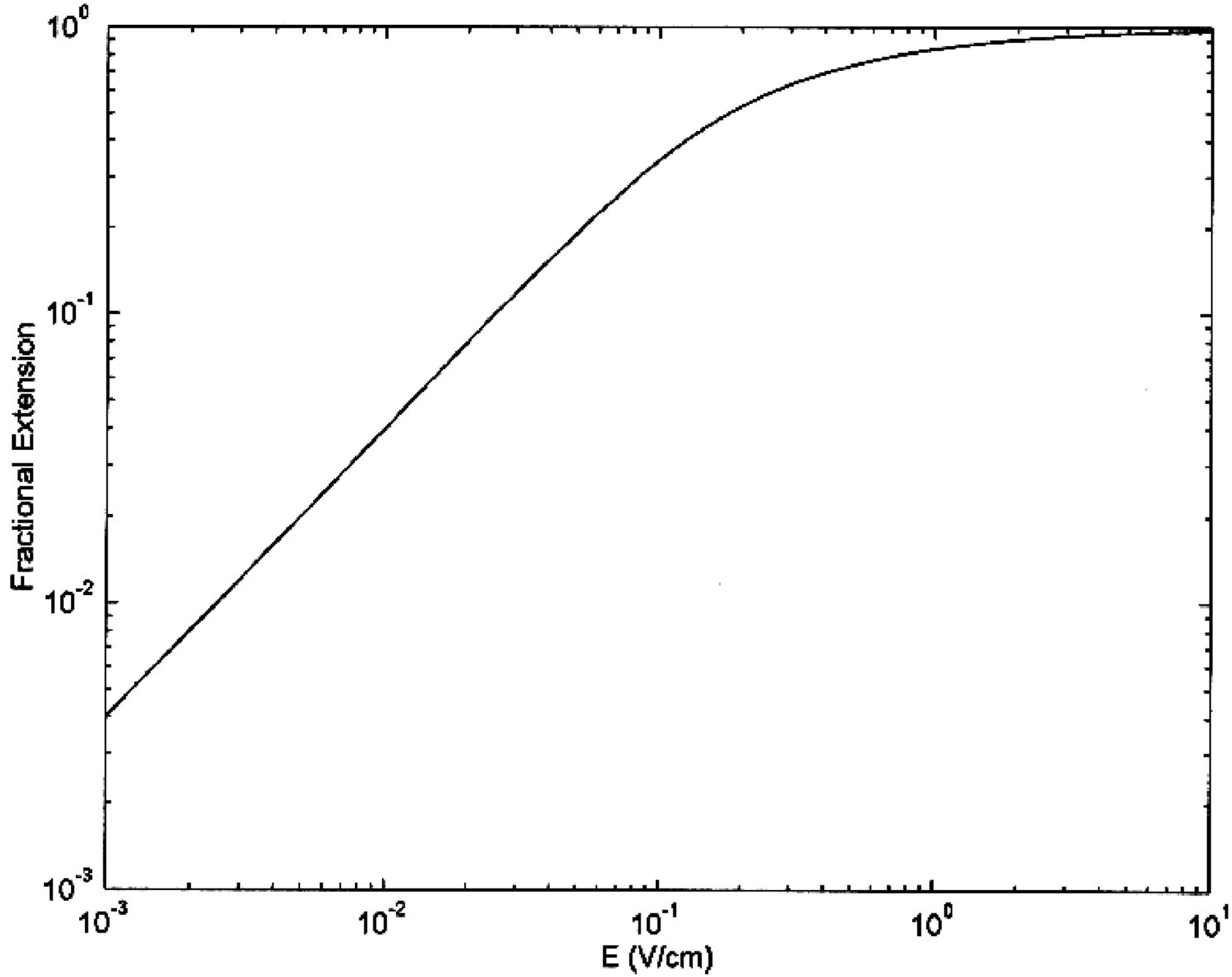
Under the free-draining condition the ELONGATION of the polymer is particularly severe because the simple additive nature of the aligning forces on each persistence length segment of the polymer.

$R_z = \kappa \ln[\sinh(L/\kappa)/(L/\kappa)]$

Where L= total length of the polymer, made of N pieces of length 2p (p=persistence length of the polymer) and:

$\kappa = 2\lambda pE/k_bT$

λ = charge/length of the polyelectrolyte.

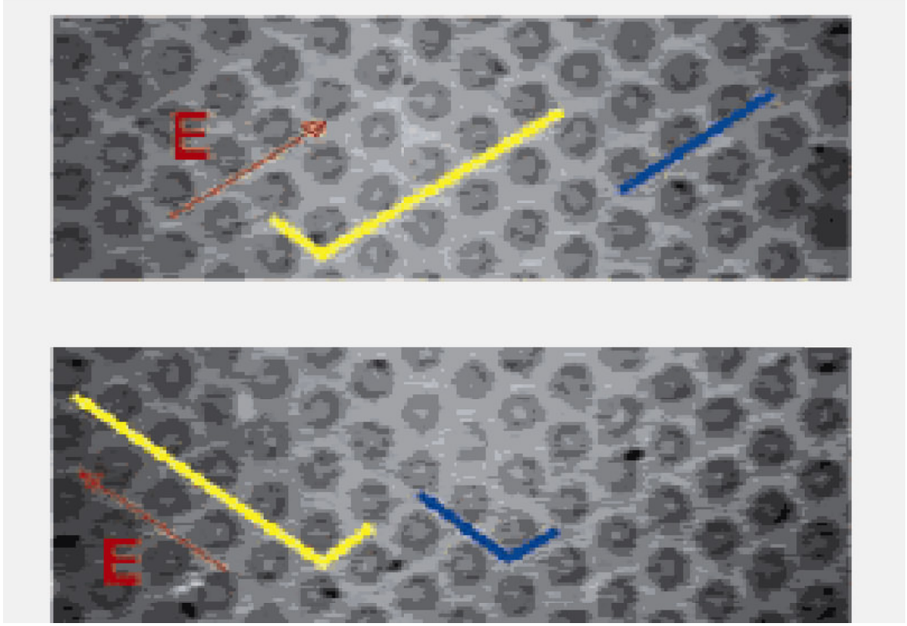


Lesson: in any confining environment, DNA is highly aligned at low fields!

This is a giant problem I have been hitting my head over for 9 years. The physics is what gets you: elongated DNA molecules in a DC field have LENGTH INDEPEN-DENT mobilities, for very elegant physics reasons maybe some of you can figure out.

I think we have found a way around this, it works for LONG DNA molecules, the question is: will it scale to nano?

The technology with best promise for ultra-long DNA fractionation is probably hex arrays coupled with crossed fields.



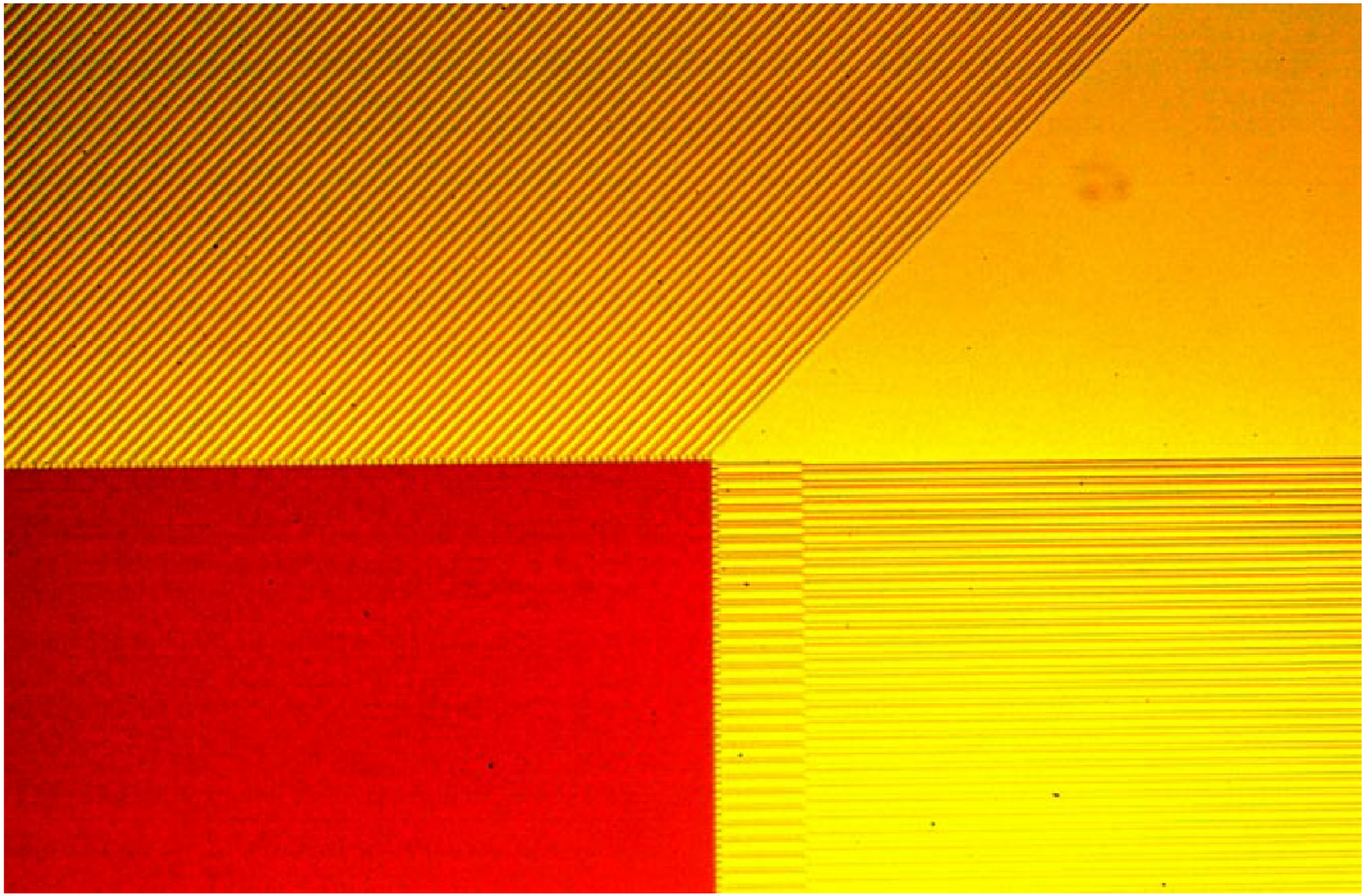
movies:demo

hex

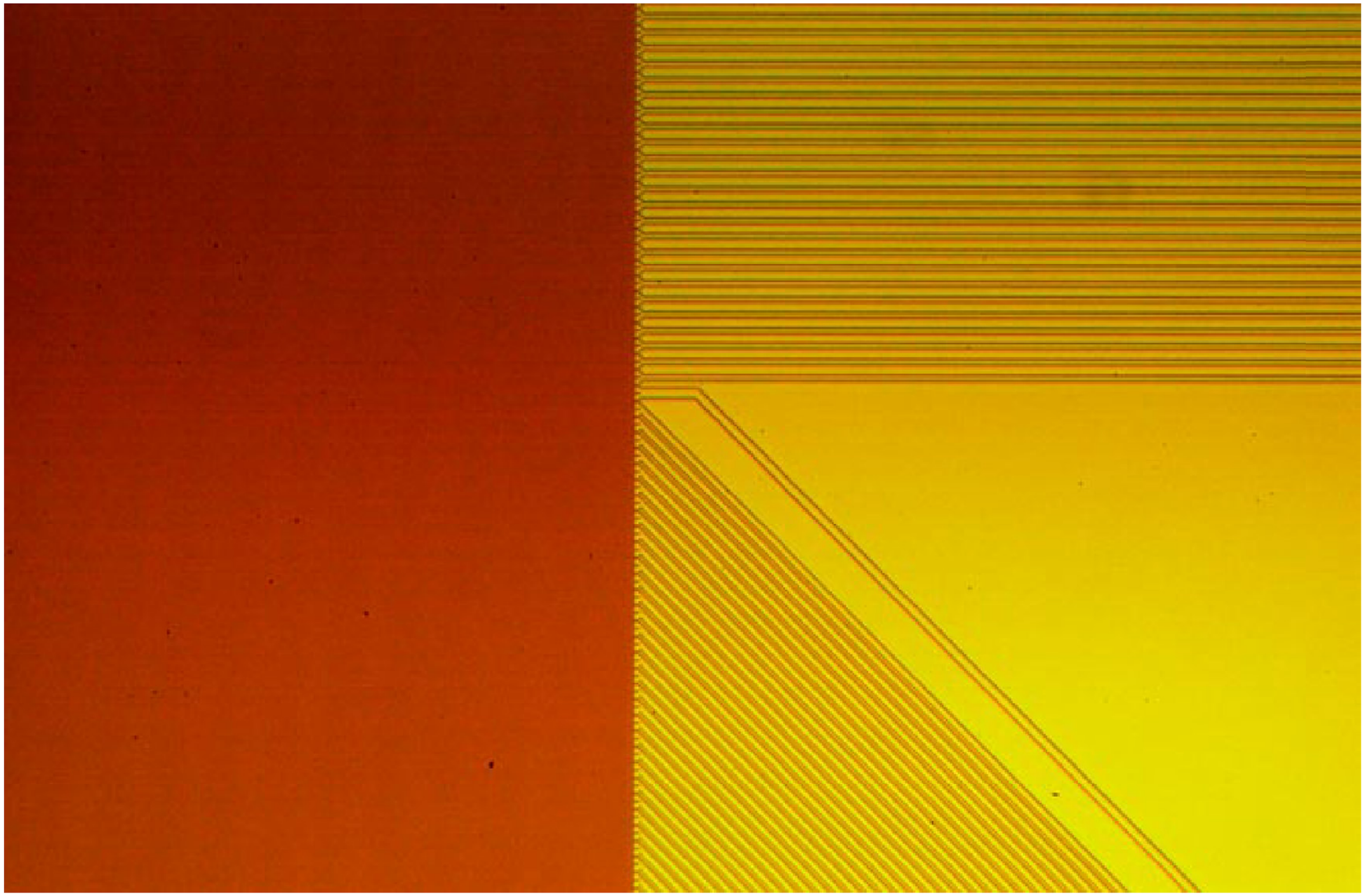
The key to this idea is the elongation of the DNA molecule that occurs in relatively high fields in a post array.

movies showing the basic physics but
hidding in the refuge of statistical analysis

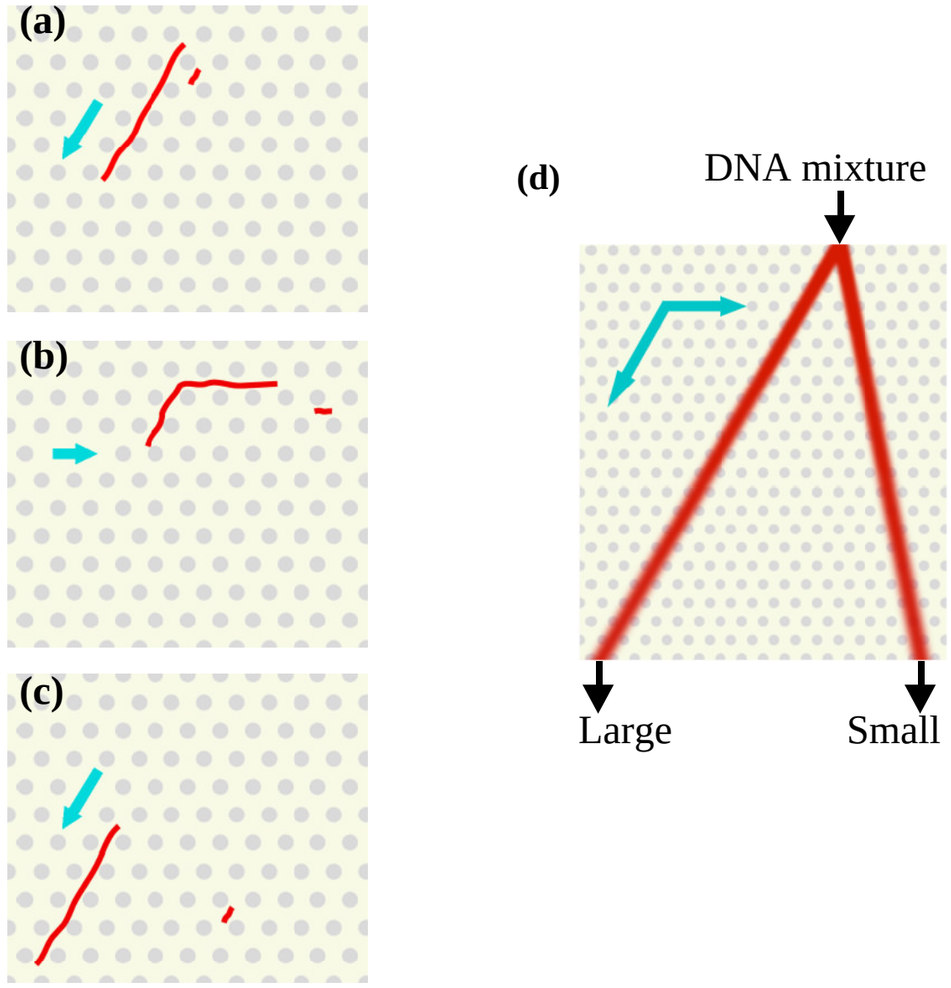
Richard Huang understood better than I did the lesson's of the Navier-Stokes equation at low Re: the current flow is like a resistive net, and can be modeled as such. The large open areas of our chips are like low impedance shunts, and injection of current through a small tube of high impedance turns a voltage source into a current source. The walls of the chips are insulators and all flow must be parallel to the walls. To achieve a uniform flow field you need what I call an N-port injector, a curtain of current sources flowing across the space into a current of sinks.



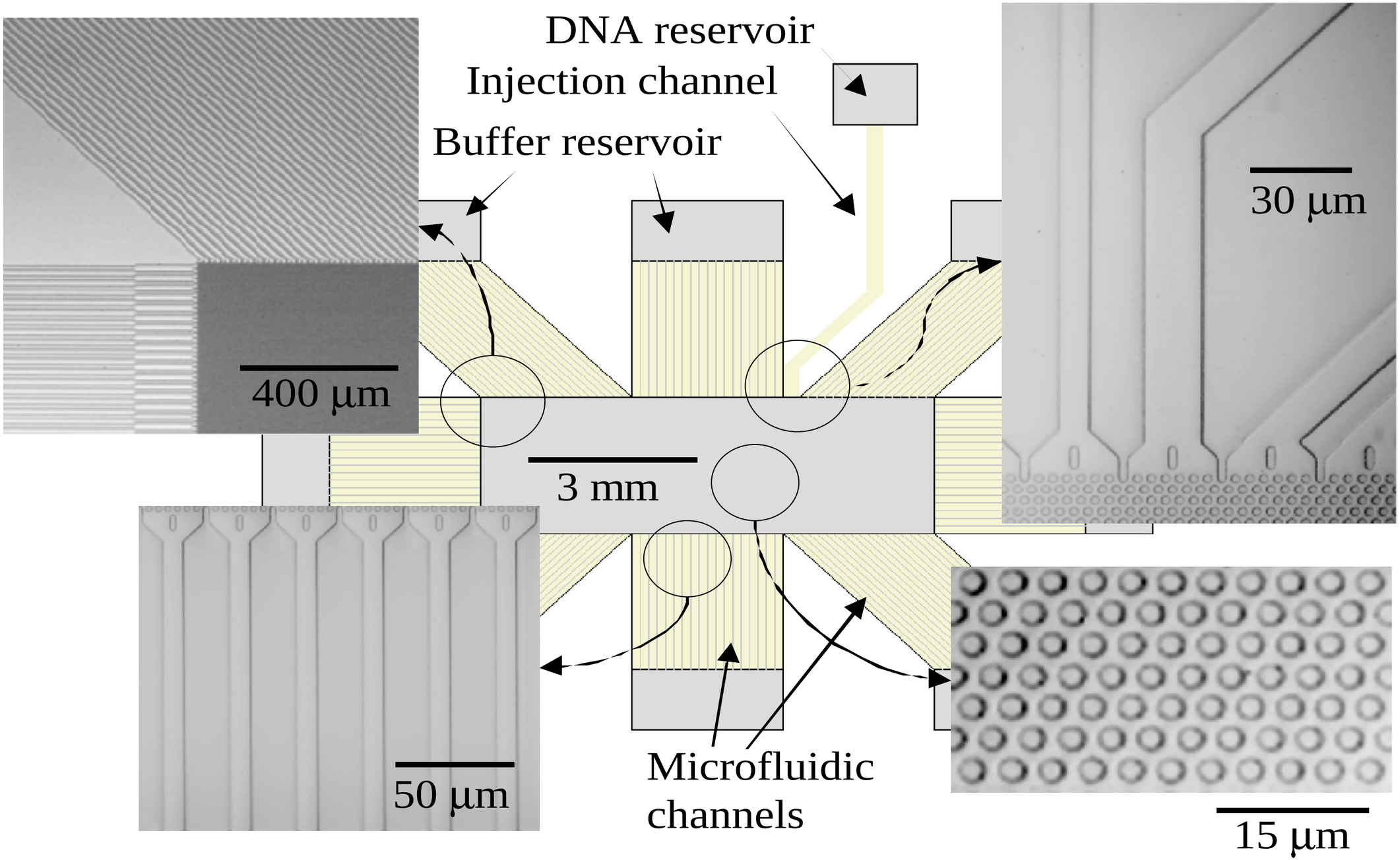
One small corner of a hex array, with 4 micron channels feeding the current flow.



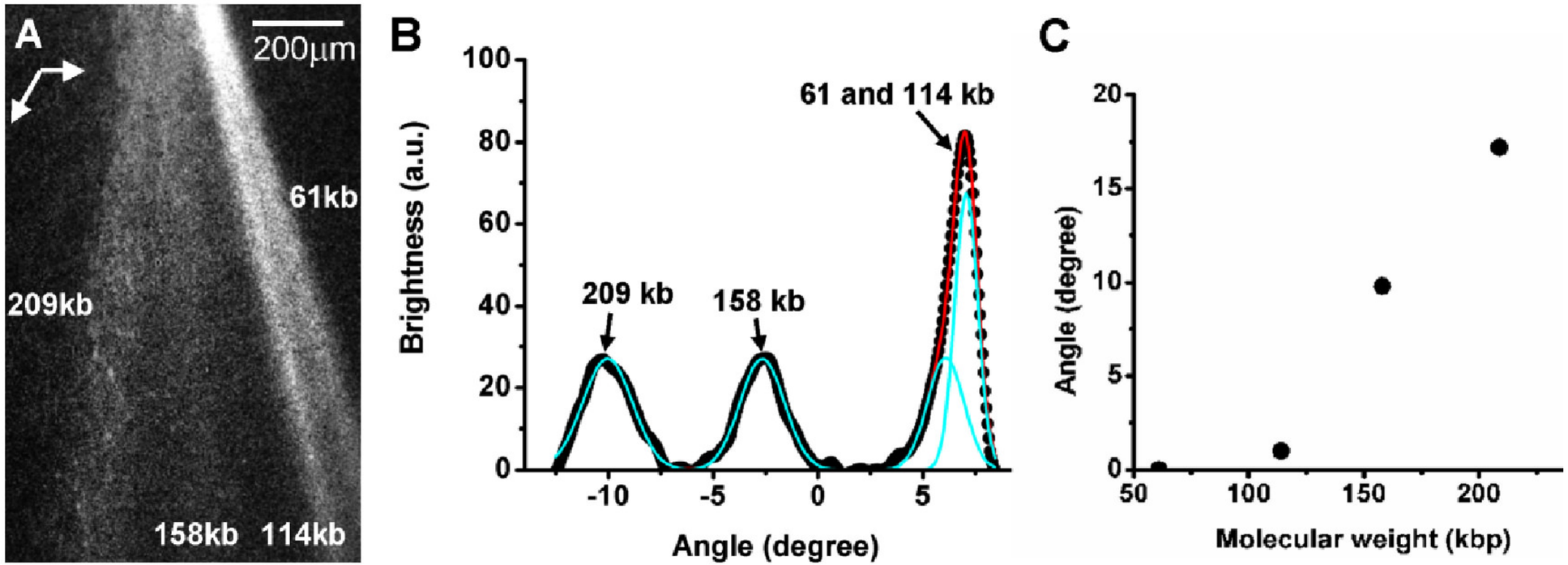
One channel is special: it brings in the DNA. This is the sample injection port. This flows across the chip like a jet stream (with no dips!)



The "DNA Prism": continuous injection and fractionation of genomic length DNA through asymmetric pulsing in a current sheet hex array.


DNA reservoir
Injection channel
Buffer reservoir
400 µm
3 mm
30 µm
50 µm
Microfluidic
channels
15 µm

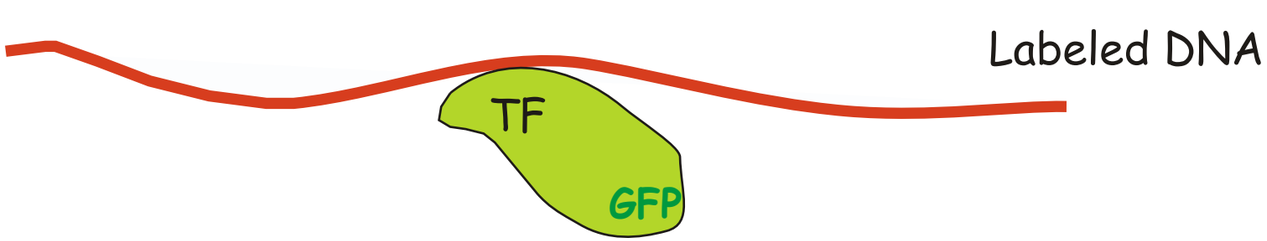
This movie is better than *BLUE CRUSH*



This does in 10 seconds what normally takes
5 hours, so that's great. But, it isn't good enough.

BUT....will this scale to nanoposts for sequencing resolution? I'm sure we need to make the posts with a nanospacing to ensure elongation, I'm not sure what happens as we scale down.

The next step is the scanning of the genomic DNA for the pattern of occupied control sites using the technology of fusion proteins between GFP and selected control proteins. Too bad Shirley T. went to the DARK SIDE.



DNA with a GFP fusion transcription factor (TF)

<> Stretch fluorescently labeled DNA with bound TF
<> Excite fluorophores using nearfield excitation
<> Detect single transcription factor molecules

Length scales and requirements

- Human chromosome #1: 279Mbp <-> 95mm
- Typical gene: 1000bp <-> 340nm
- Typical TF binding site: 20-40bp <-> 7..14nm

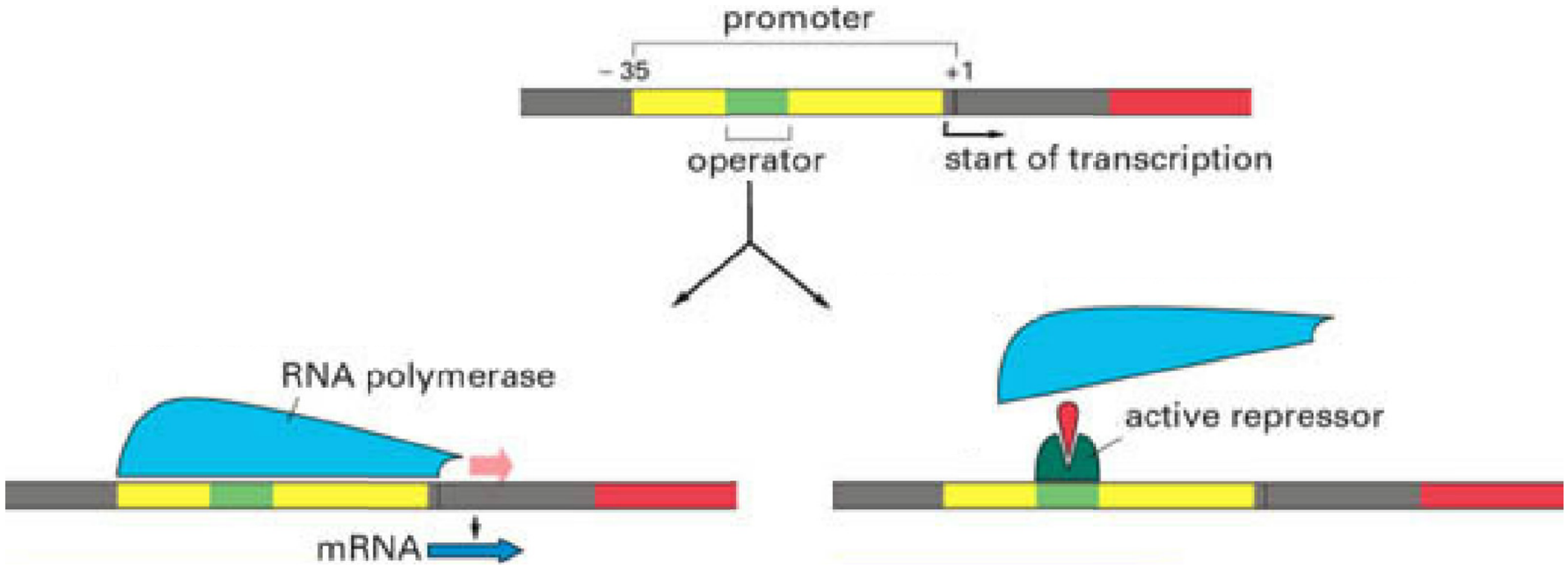
Assign each transcription factor to a specific gene or a specific binding site:
required resolution 340nm to 7nm -> Use **nearfield** excitation.

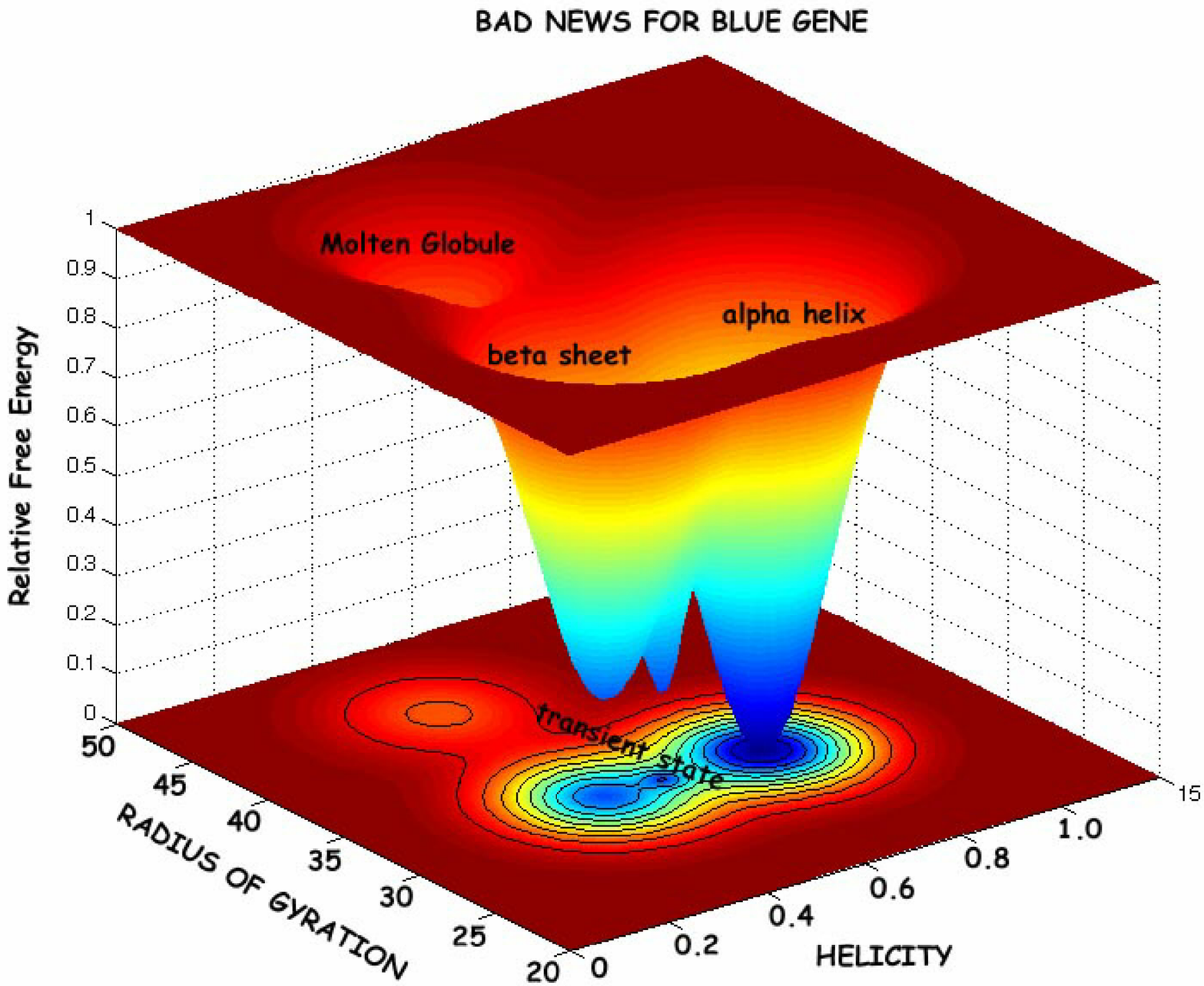
Length units: "base pairs" not "nm"
-> stretch DNA **uniformly**

Model System

Our model system is the lac operon, which controls the lactose metabolism of E. Coli. The lac repressor (lacI) is a transcription factor that binds to the DNA and prevents the transcription of the operon. The Tilghman lab provides us with lacI-GFP (lac repressor fused with GFP) and model DNA with multiple sites of the lac operator (binding site for the lacI-GFP repressor protein). Hungarian Physicist Leo Szilard.

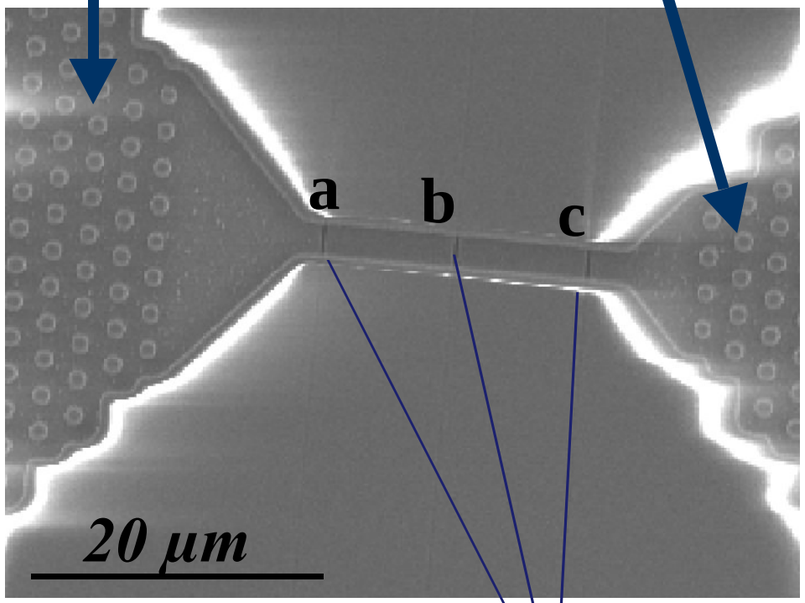
How a repressor works



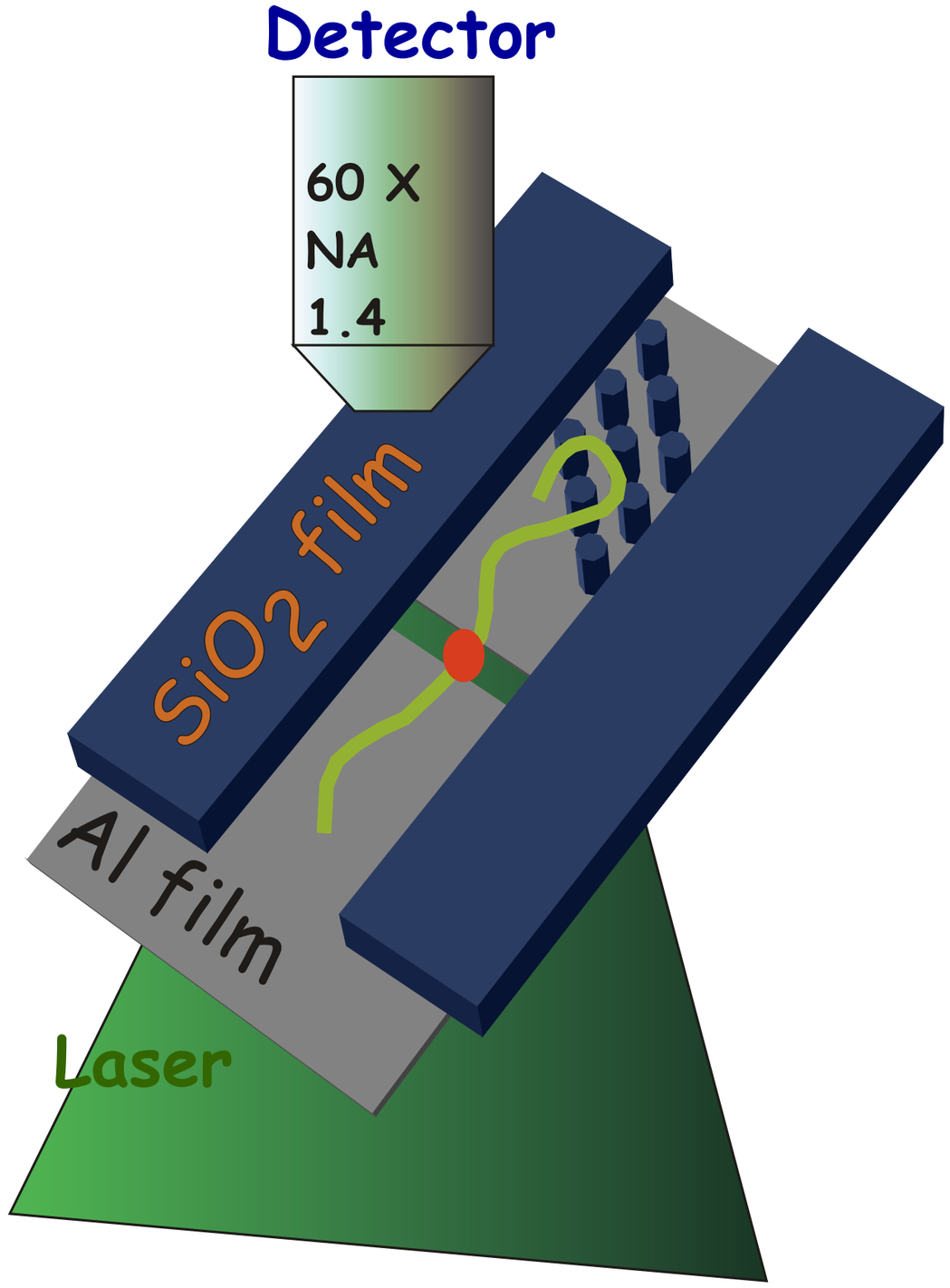

BAD NEWS FOR BLUE GENE
Molten Globule
alpha helix
beta sheet
transient state
Relative Free Energy
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
50
45
40
35
30
25
20
RADIUS OF GYRATION
0
0.2
0.4
0.6
0.8
1.0
HELICITY
15

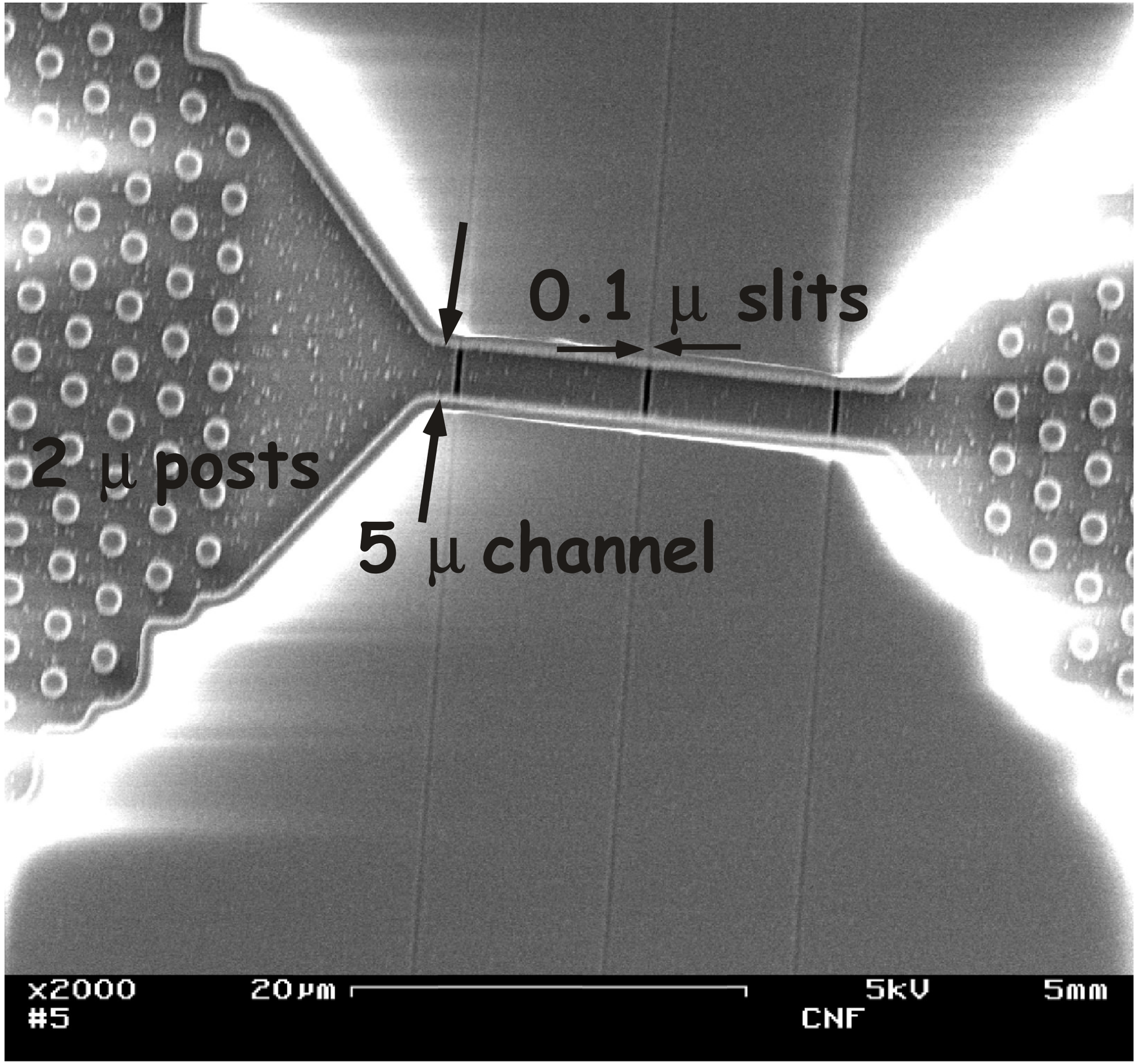
DEVICE

Posts for stretching the DNA (UV litho)



Slits for light width: 40nm-1μm (e- litho) beam

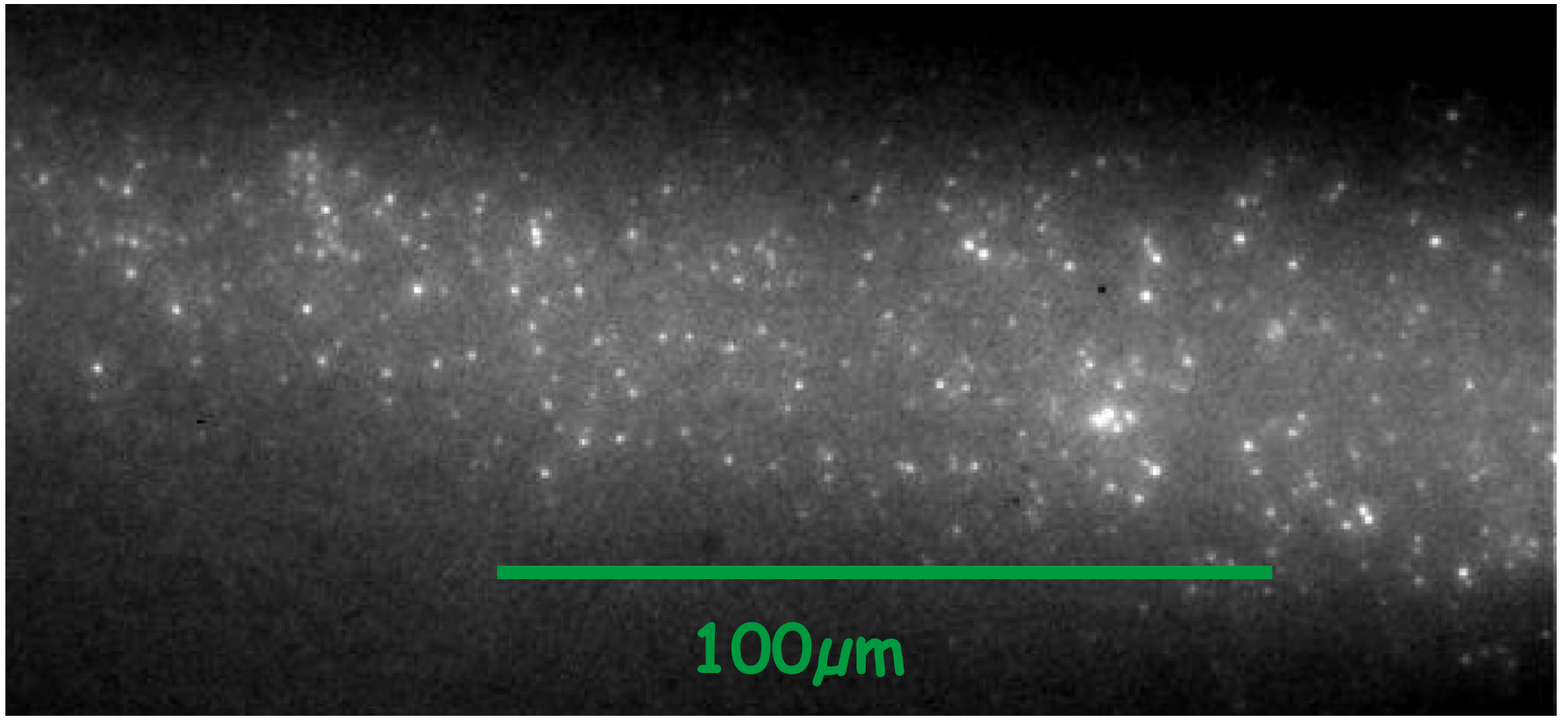



0.1 μ slits
2 μ posts
5 μ channel
x2000
20μm
5kV
5mm
#5
CNF

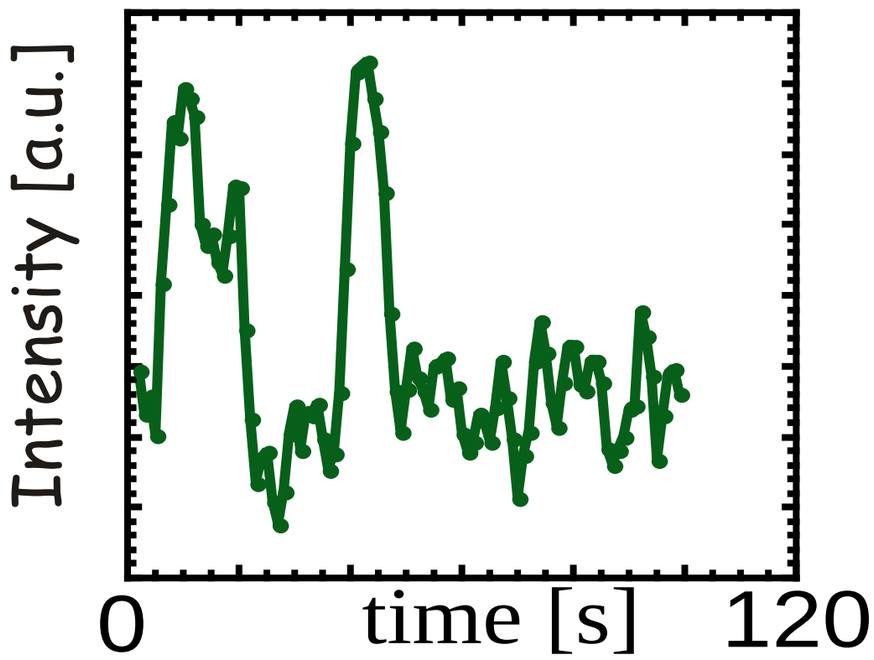
MOVIE SHOWING THE BASIC DYNAMICS OF DNA IN A CHANNEL

Single lacI-GFP protein molecules

Concentration 20nM (10 GFP per μm^3)



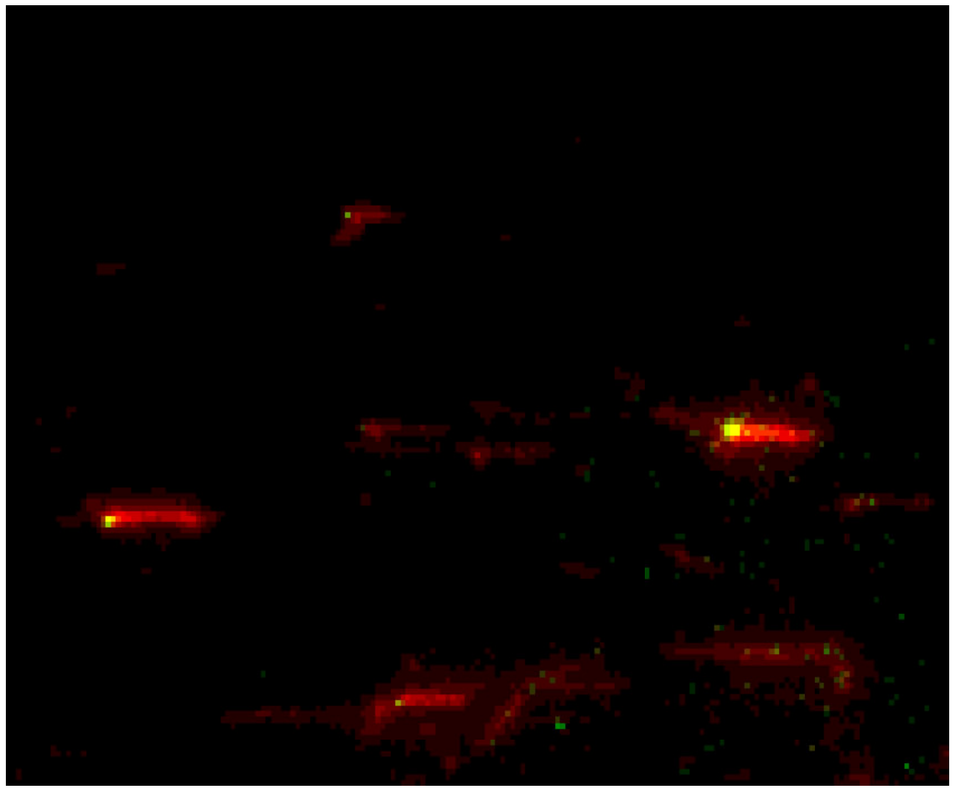
The fluorescence intensity from one GFP is followed in real time. Characteristic for GFP is that the fluorescence emission switches on and off.



Repressor binding to DNA

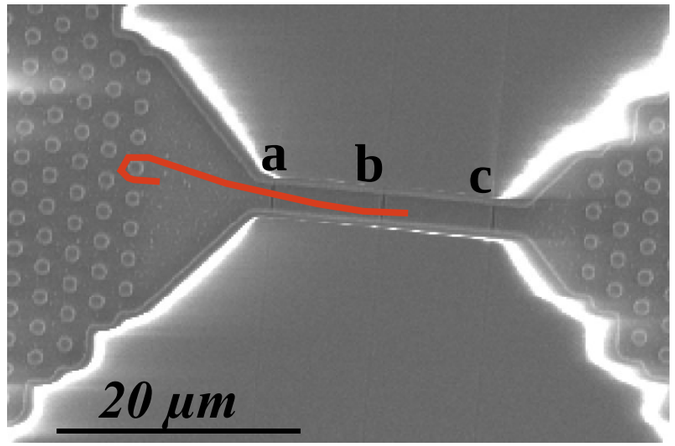
(preliminary results)

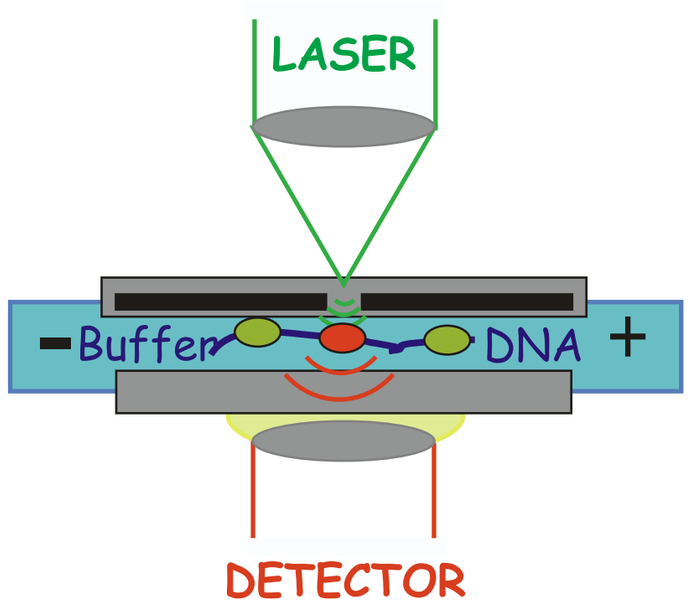
Lac repressor (lacI GFP) is allowed to bind to multiple lacO sites inserted in a λ DNA. The DNA is then labeled with a red intercalating dye (BOBO3). To visualize the GFP, blue excitation light is used. Next, for the red BOBO3 yellow excitation light is used. The two resulting pictures are combined into the following false color rendering. The green spots correspond to GFP and the red correspond to the DNA.



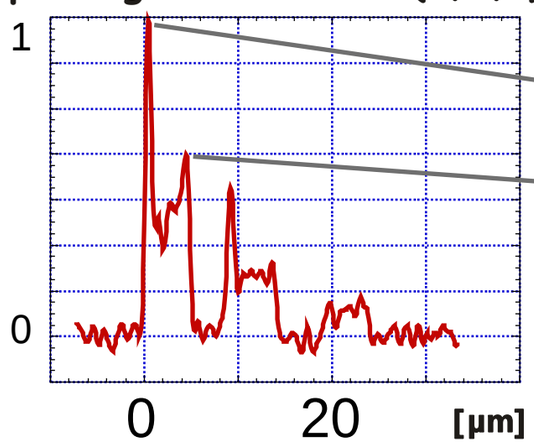
20μm

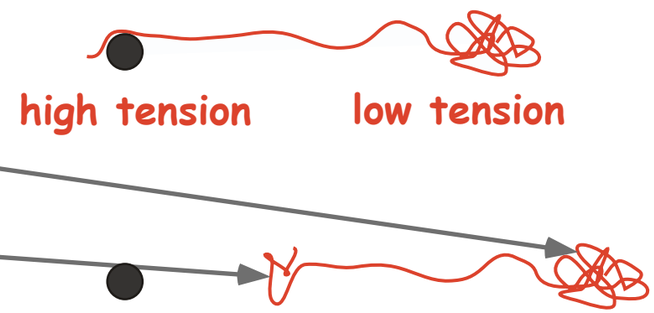
MOVIE SHOWING THE TRANSIENT LIFETIME OF SINGLE MOLECULES, YFP-CI FUSION PROTEINS





Fluorescence signal from DNA passing 100nm-slits (a,b,c)





After release from the post the DNA relaxes from both ends resulting in a disordered head and tail.

There is a company, US genomics (read about them in the New York Times) that is exploring this basic technology. They do the engineering right.

Poor stretching of DNA

The ends are clearly disordered in the DNA since the DNA is moving in channels much larger than the persistence length of the DNA. By forcing the DNA into channels that are on the order of the persistence length (50nm), the DNA is forced into a uniformly stretched confirmation.

Poor Optical Resolution (200nm)

The nearfield intensity and resolution falls off rapidly with distance from the nearfield slits. Smaller channels force the DNA to close proximity of the slits increasing the intensity and the resolution of the nearfield excitation.

The real challenge now is to make ARRAYS of nanochannels to bring a number of genomic length DNA molecules, stretch them to their full length and scan them for the protein occupation sites...and with all respect to Harold "E-beam litho" Craighead, we would really like to avoid e-beam litho because it is relatively slow and expensive.

We've had some good luck lately using Steve Chou's beautiful imprinting technology, which allows you to transfer from a master mold into a substrate at the nanoscale.

We are hard (really hard) at work on the Mark II chip design:

(1) Nanoimprinted arrays

(2) Evanescent wave excitation using critical angle excitation to get true single molecule detection.

(3) Nanofabrication of channels as well as slits, using focused ion beam (FIB) technology and nano-imprinting.

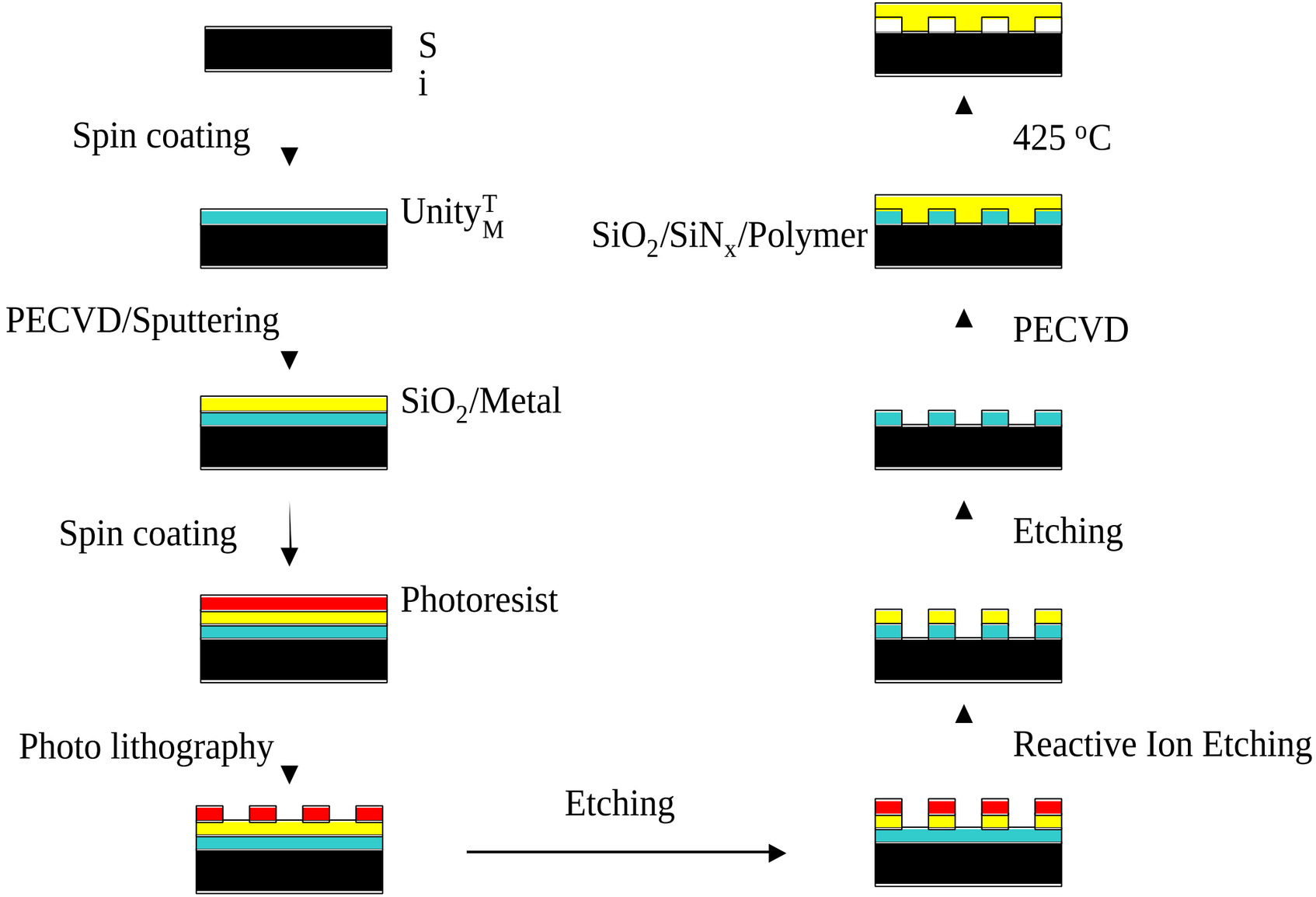
(4) "Unzipping" polymer technology to create buried internal channels at the nanoscale (with B.F. Goodrich, polynorbornene)

(4) Sealed nanostructures using unzipped polymers.

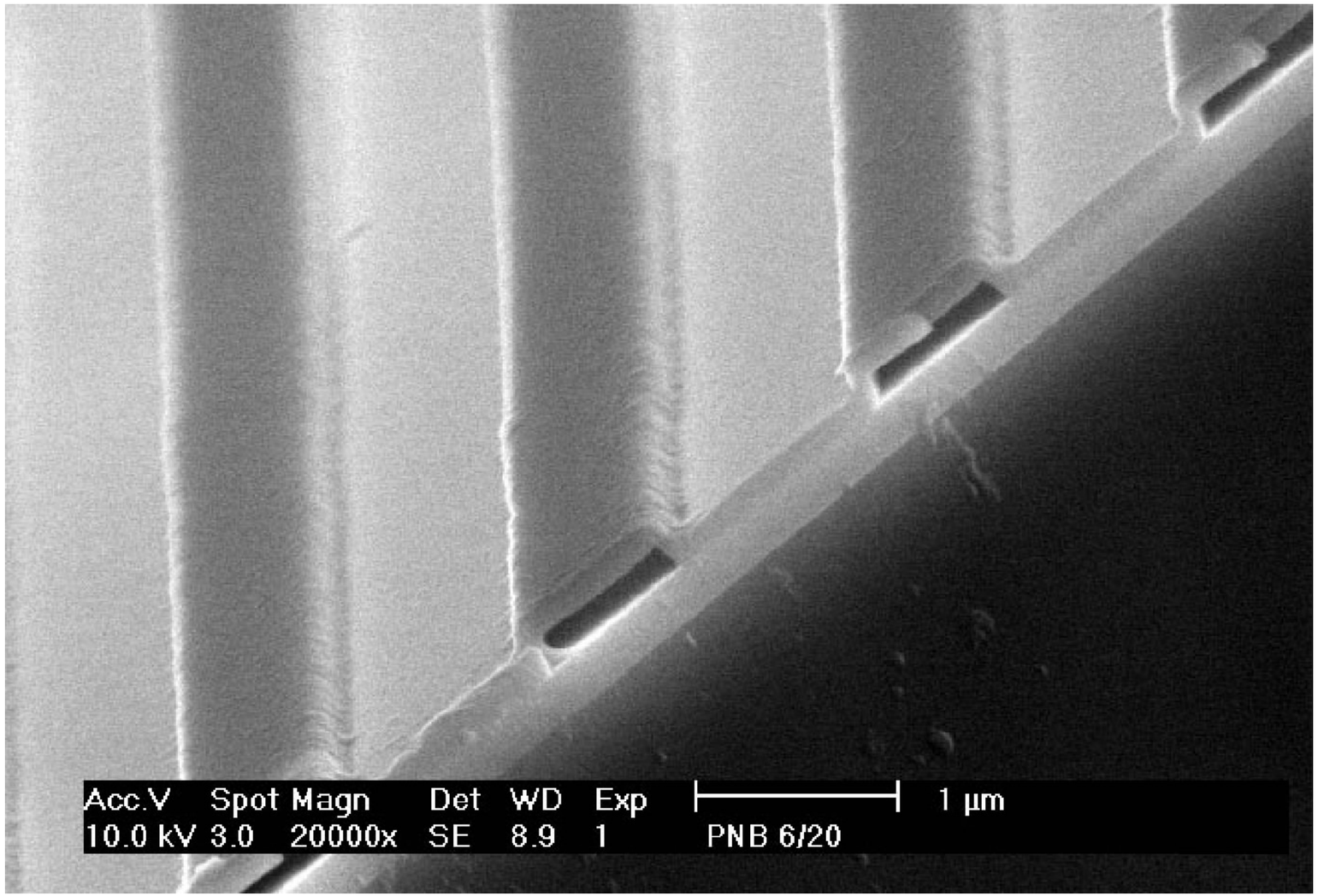
"Unzipping" polymers are polymers which undergo a phase change directly from solid to gas. A layer of such a polymer can be nanomachined to form structures, chemical vapor deposition techniques can be used to cover the nanostructures with quartz (for example), then the polymer can be unzipped, leaving behind sealed nanostructures.

We have successfully made these structures, it is a difficult process requiring ultra-high vacuum and careful temperature control. Collaboration with B.F. Goodrich.

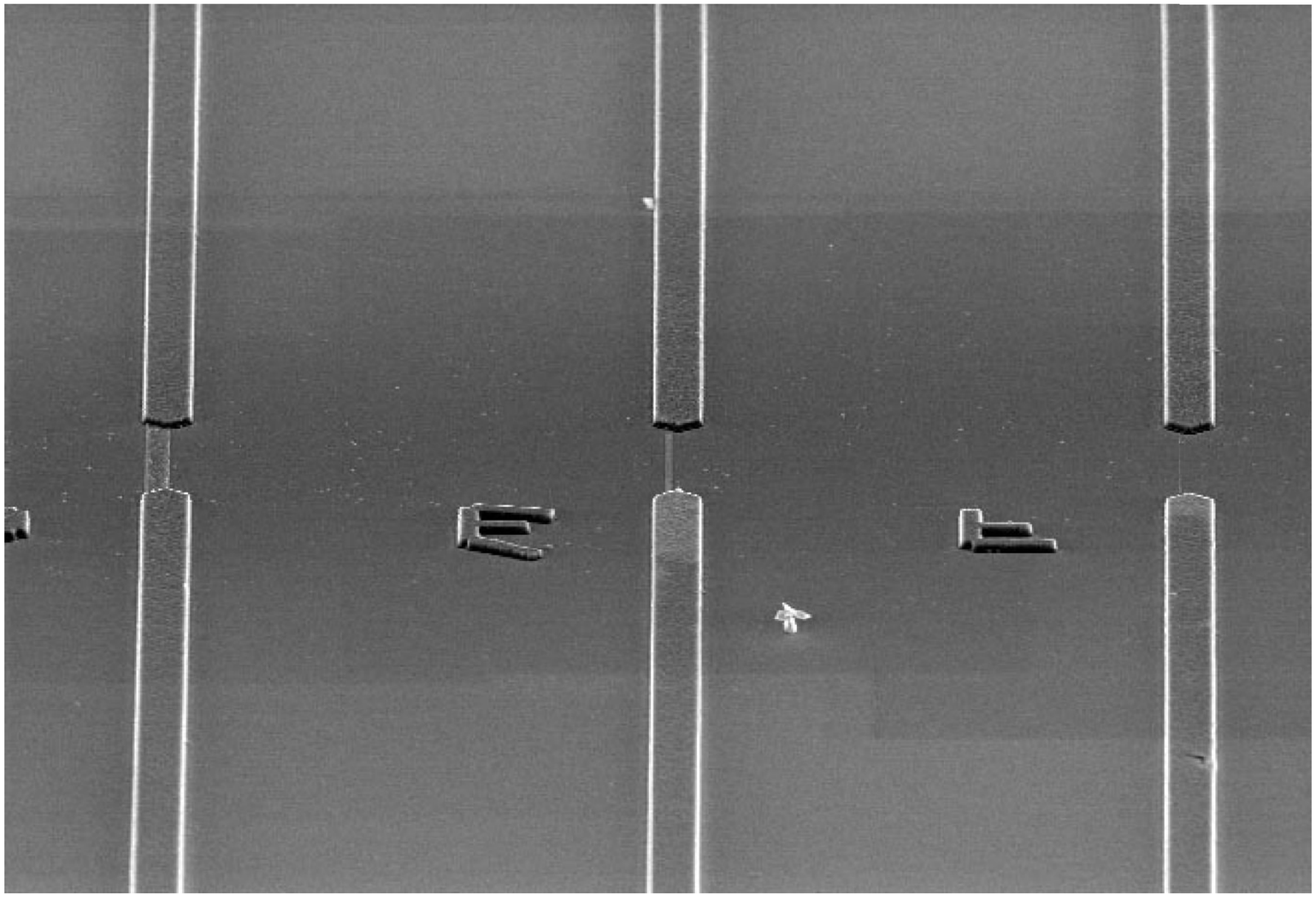
Example of Channel Formation



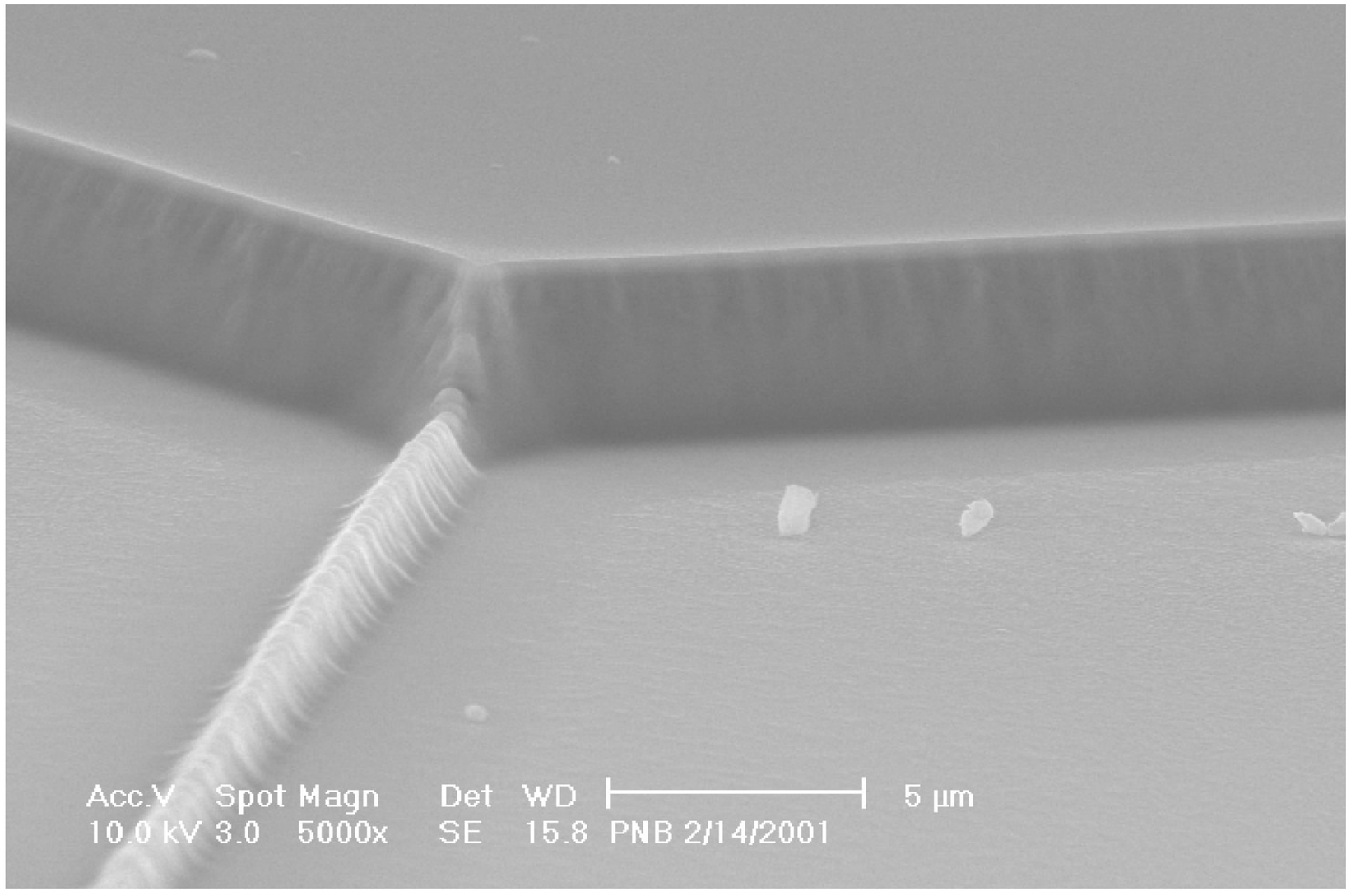
It still is a rather complex process



clean channels 100 nm deep

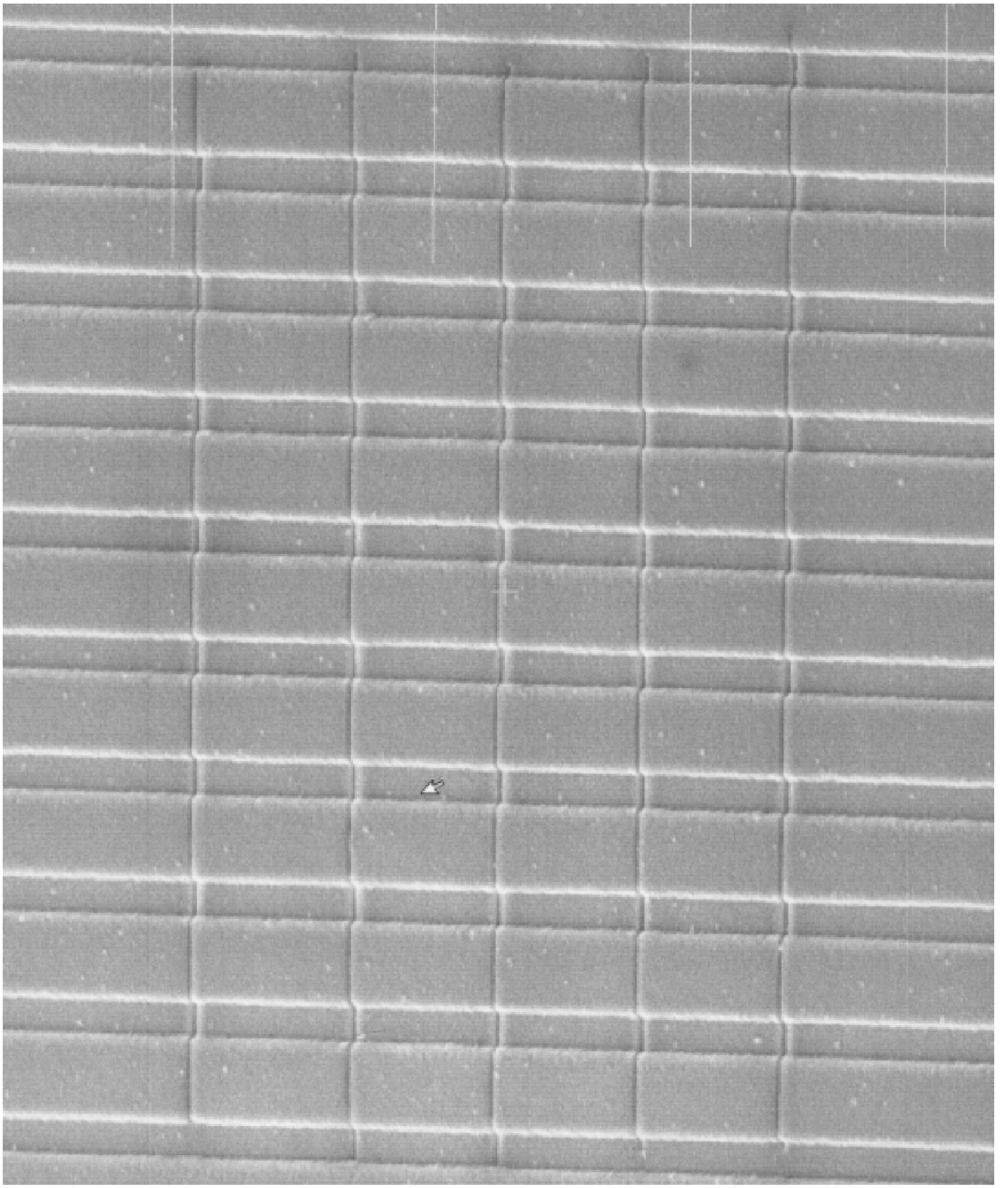


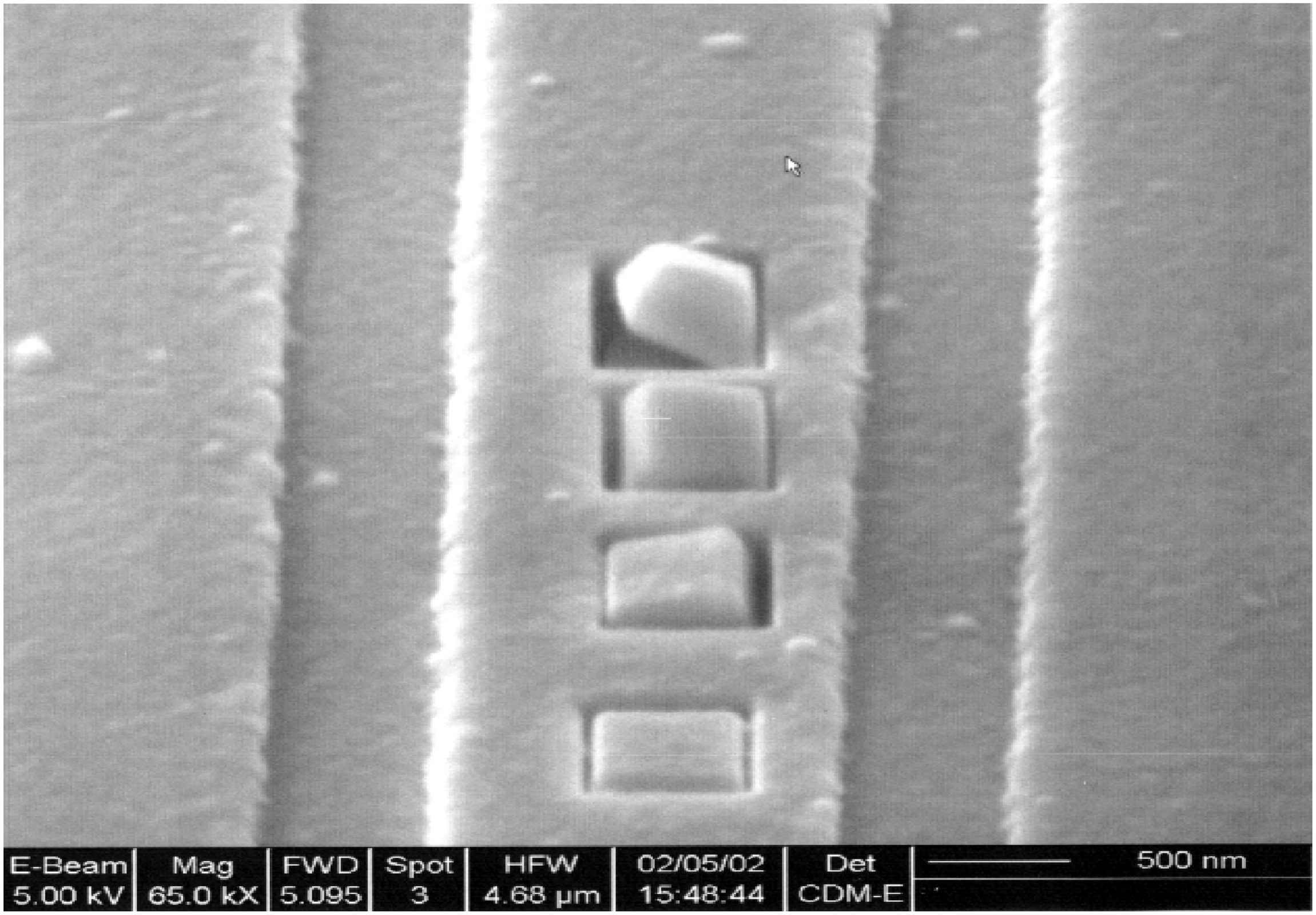
3 channels: 5, 0.5, 0.05 microns wide using e-beam lithography


Acc.V Spot Magn Det WD 5 µm
10.0 kV 3.0 5000x SE 15.8 PNB 2/14/2001

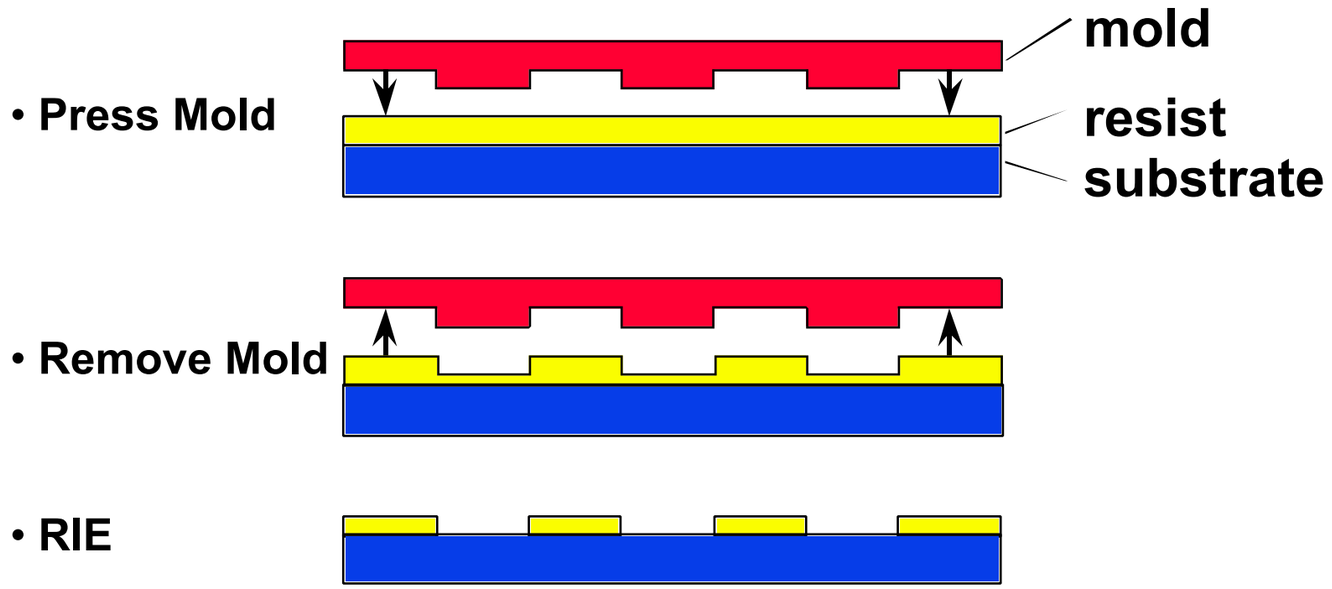
(3) FIB: Focused Ion Beam Nanomachining.

This is a rapidly growing technology which uses energetic ions, typically gallium+, to sputter away material. It is used in the semiconductor industry to "slice and dice" integrated circuits. We have done preliminary FIB on our structures (using U. Maryland FIB machines) to show that we can cut 50 nm wide slits directly into aluminum films. This eliminates many of the tedious aspects of e-beam lithography, since it is a direct write/read process.



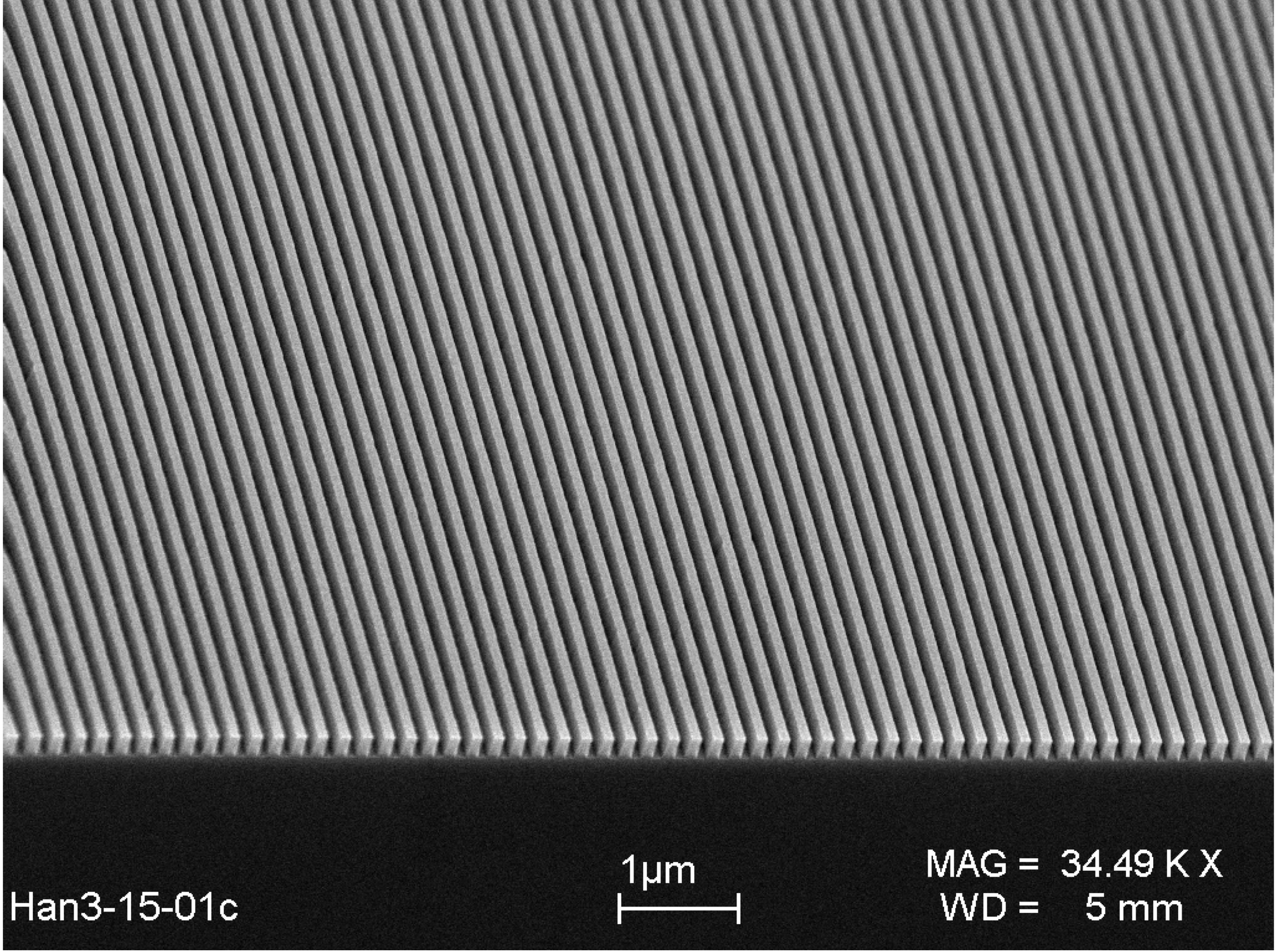

E-Beam
5.00 kV
Mag
65.0 kX
FWD
5.095
Spot
3
HFW
4.68 µm
02/05/02
15:48:44
Det
CDM-E
500 nm

Nanoimprinting



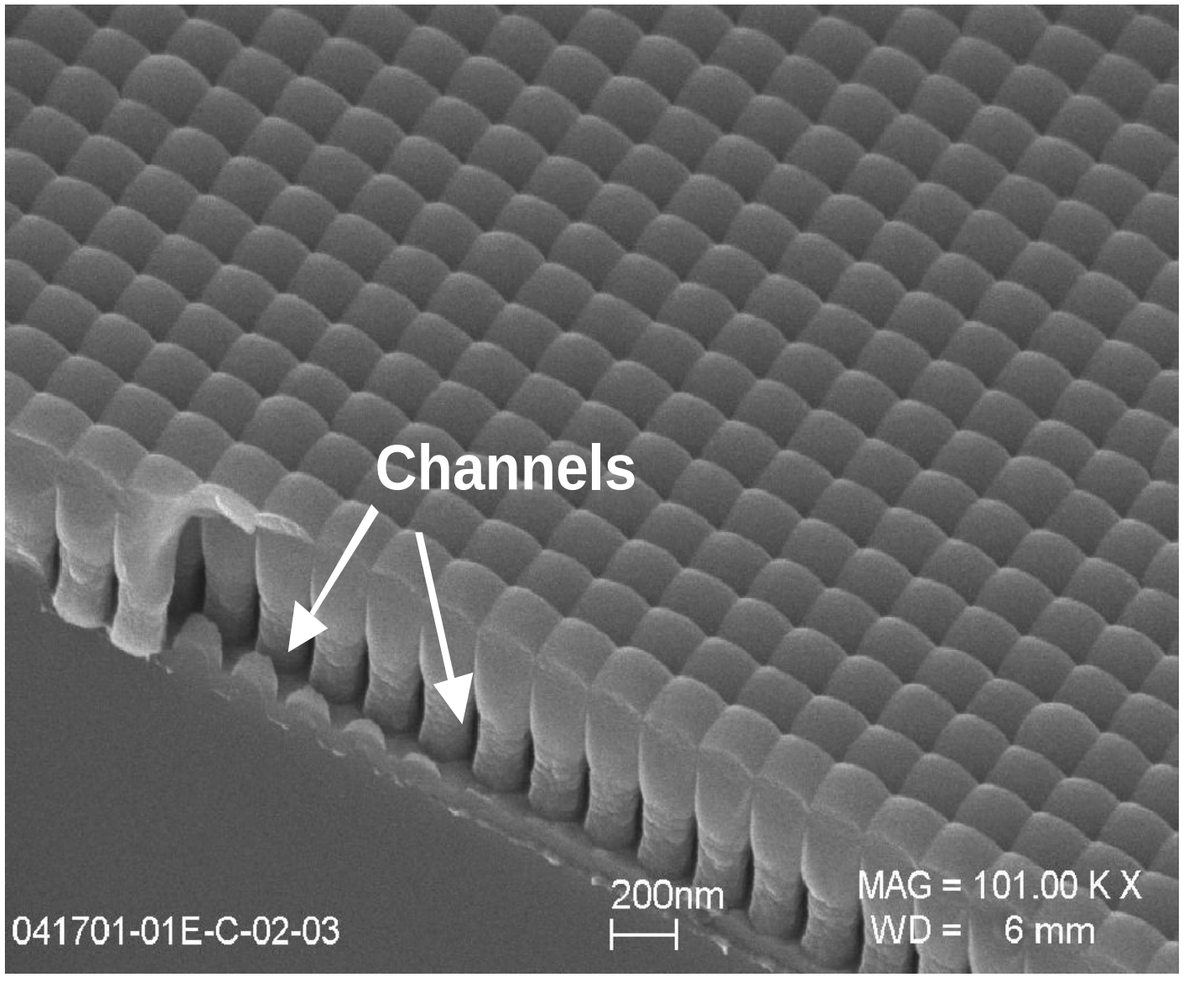
Chou, Krauss, and Renstrom, APL, Vol. 67, 3114 (1995); Science, Vol. 272, 85 (1996)

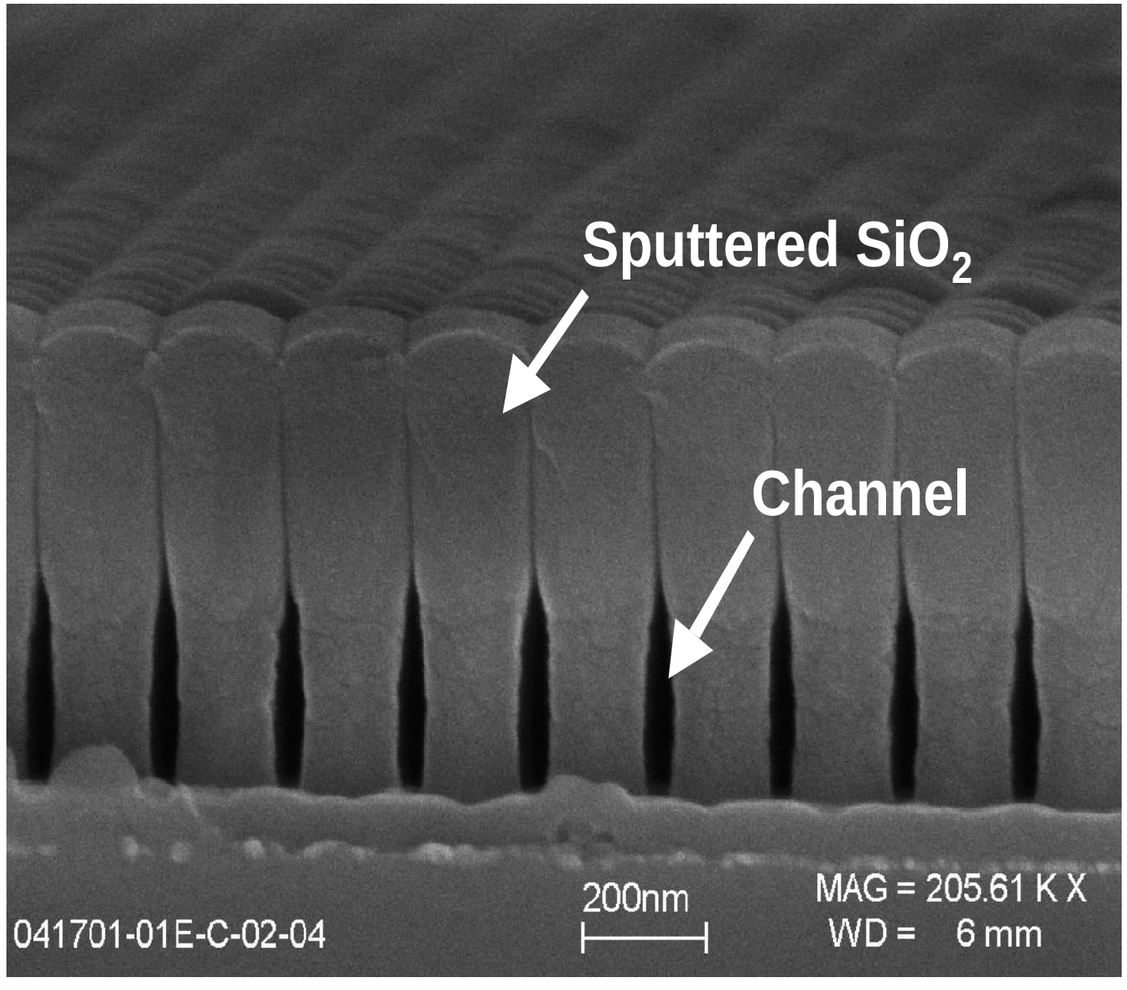
This technology allows you to transfer patterns with nanometer resolution across a full 4 inch wafer.

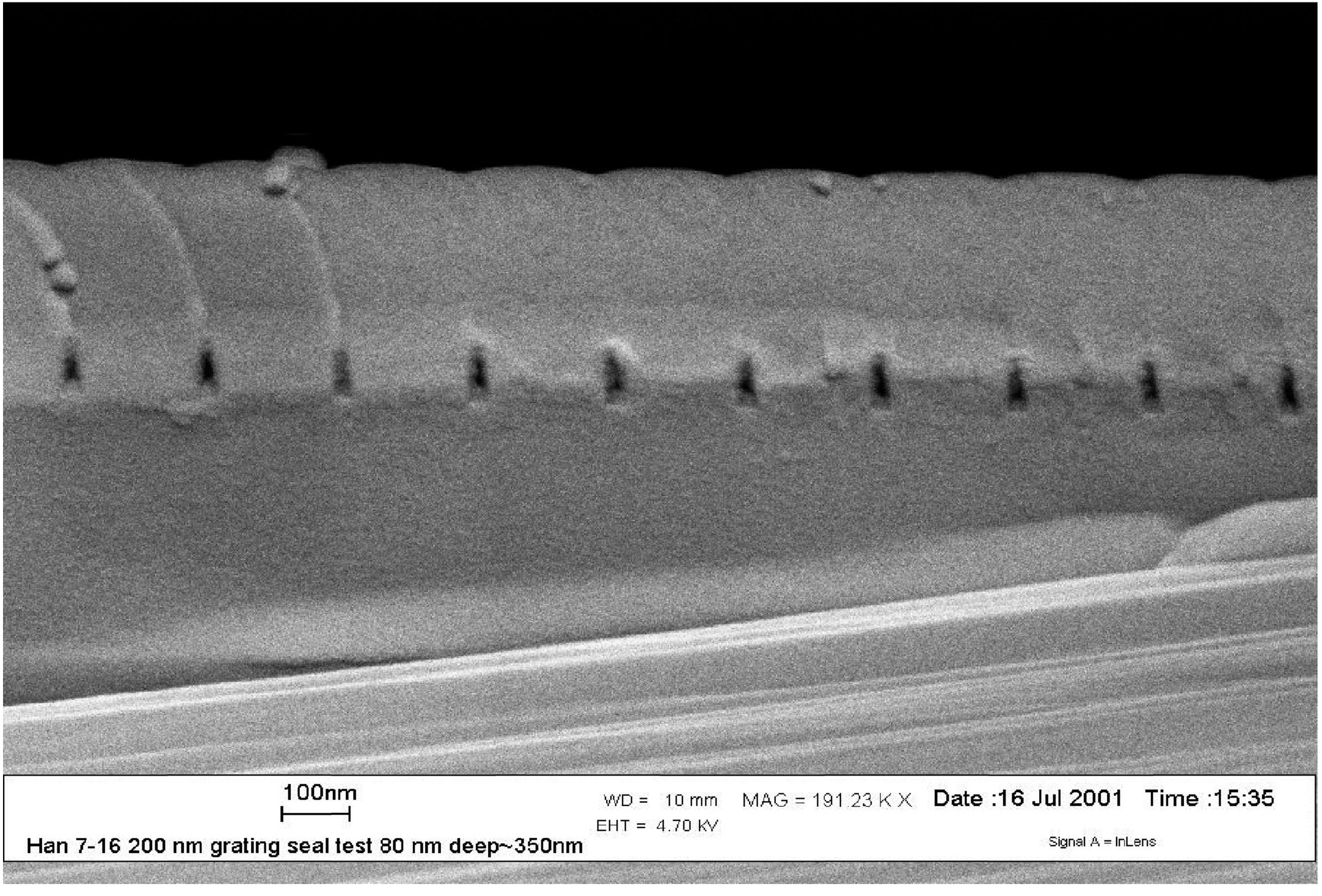

1µm
MAG = 34.49 K X
WD = 5 mm
Han3-15-01c

Sealing of these structures can be done in several ways:

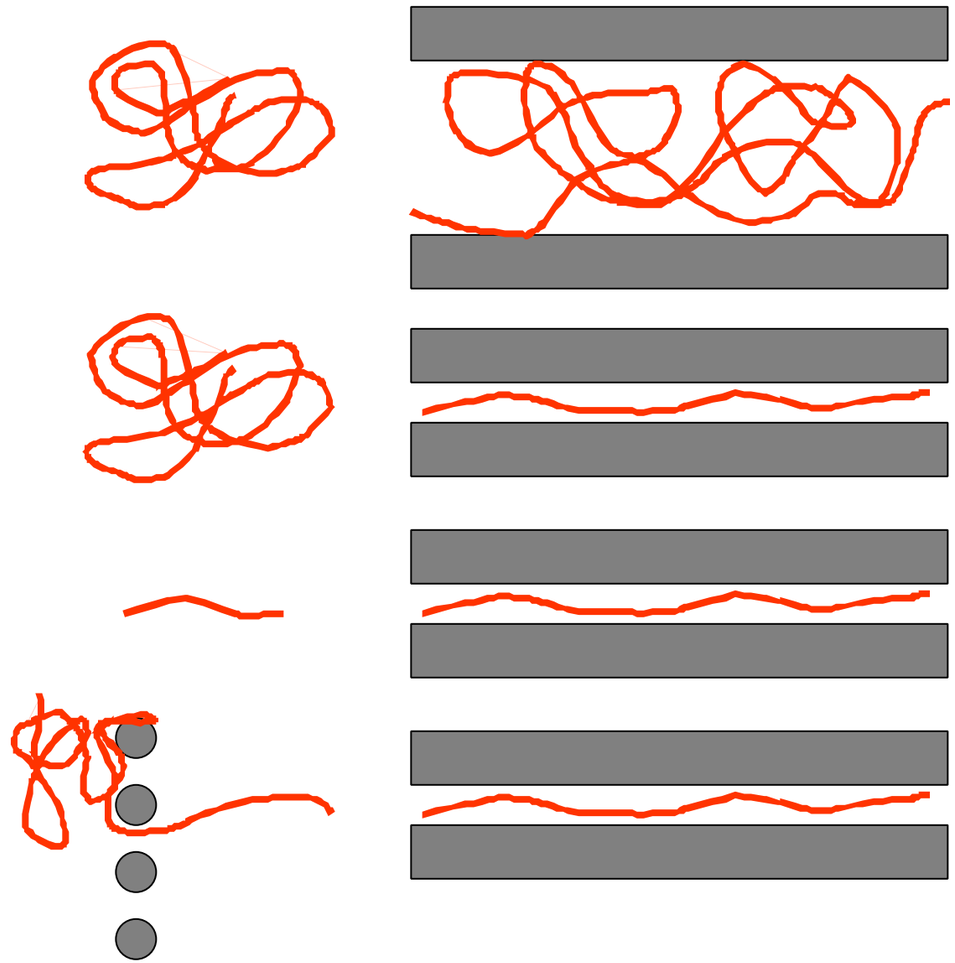
(1) silicone gaskets (not nano)
(2) oblique sputtering
(3) internally buried chan-nels


Channels
041701-01E-C-02-03
200nm
MAG = 101.00 K X
WD = 6 mm


Sputtered SiO_2
Channel
041701-01E-C-02-04
200nm
MAG = 205.61 K X
WD = 6 mm


100nm
WD = 10 mm
MAG = 191.23 K X
Date :16 Jul 2001
Time :15:35
EHT = 4.70 kV
Signal A = InLens
Han 7-16 200 nm grating seal test 80 nm deep~350nm

Large change in entropy when DNA enters nanochannel



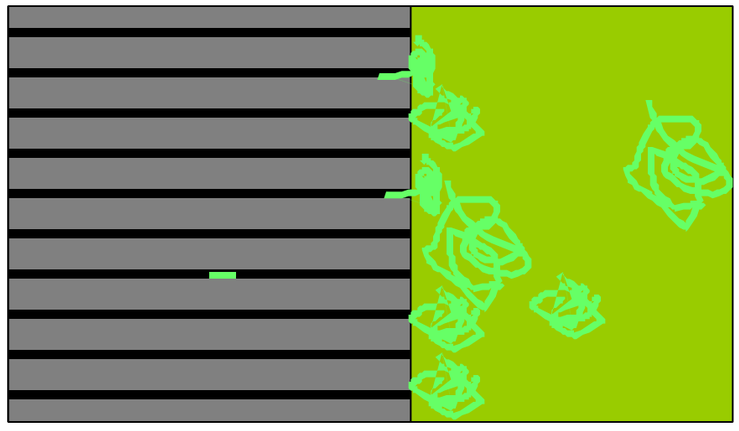
$\Delta S \sim 0$

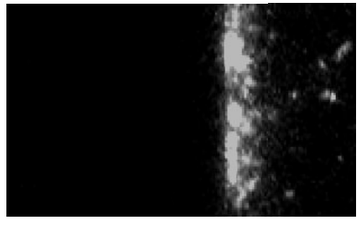
$\Delta S << 0$

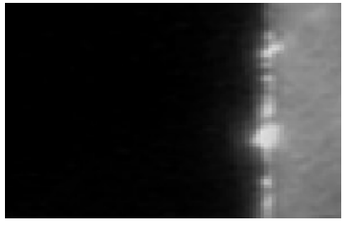
$\Delta S \sim 0$

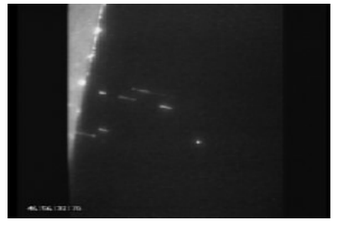
ΔS small

Long DNA Molecules Stall At edge of theNanochip

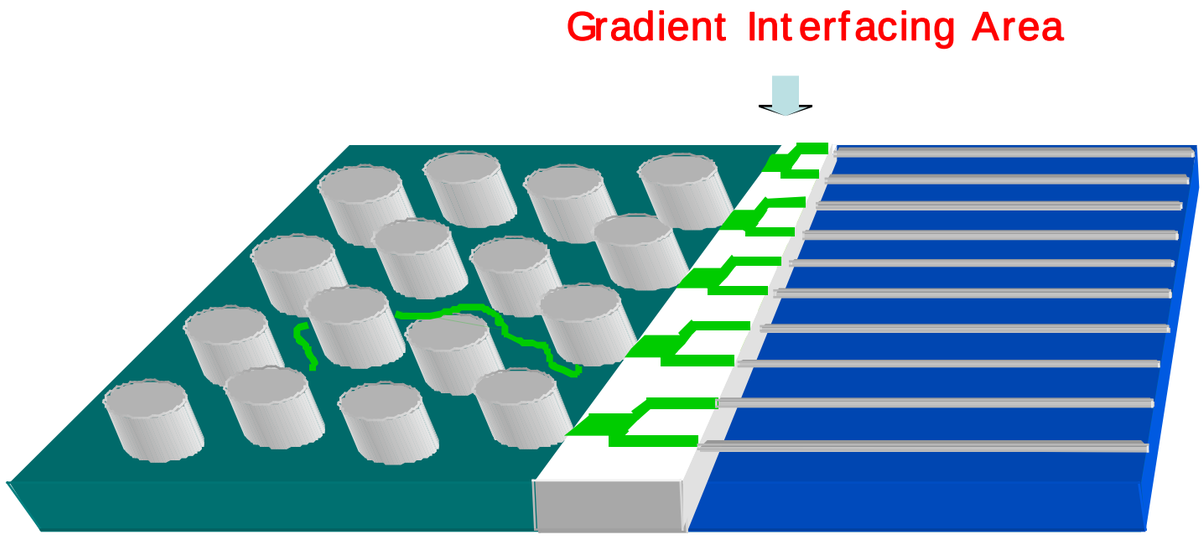


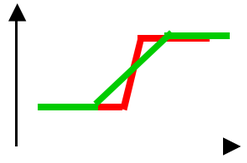






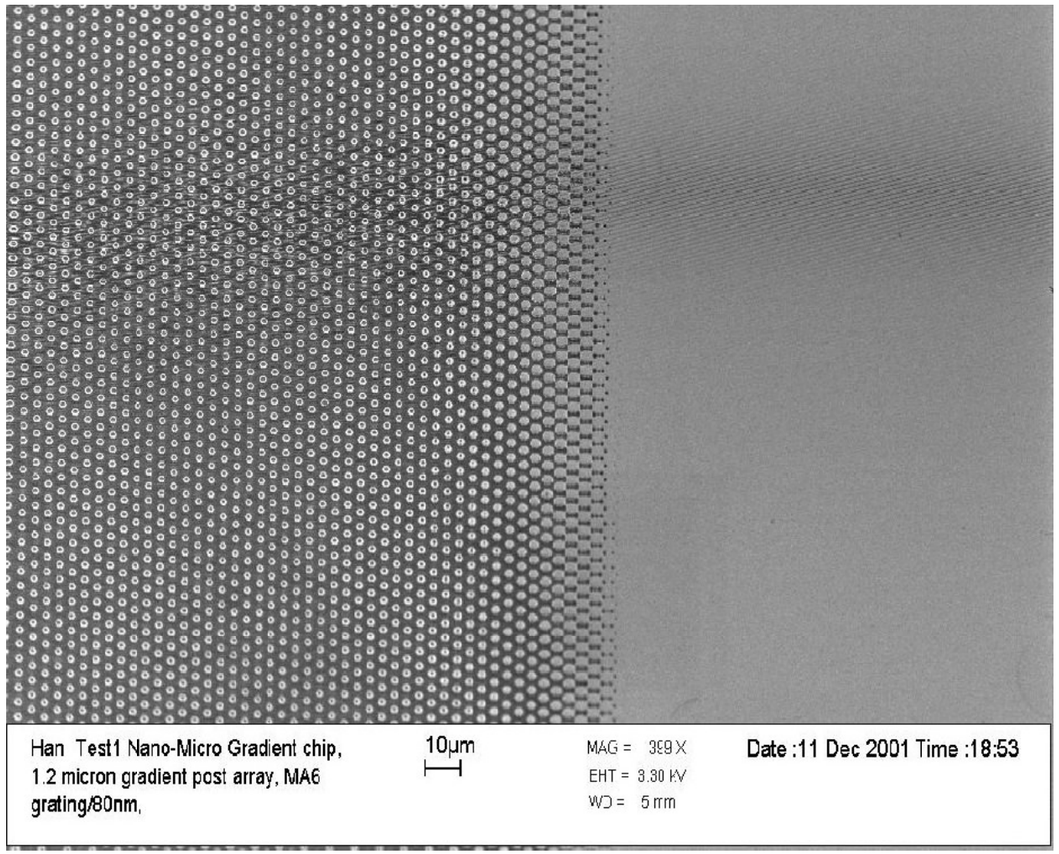
Micro/ Nano Total Analysis Systems Needs Integration of Microscale and Nanoscale Fluidics And Important Structural Component-Gradient Structures

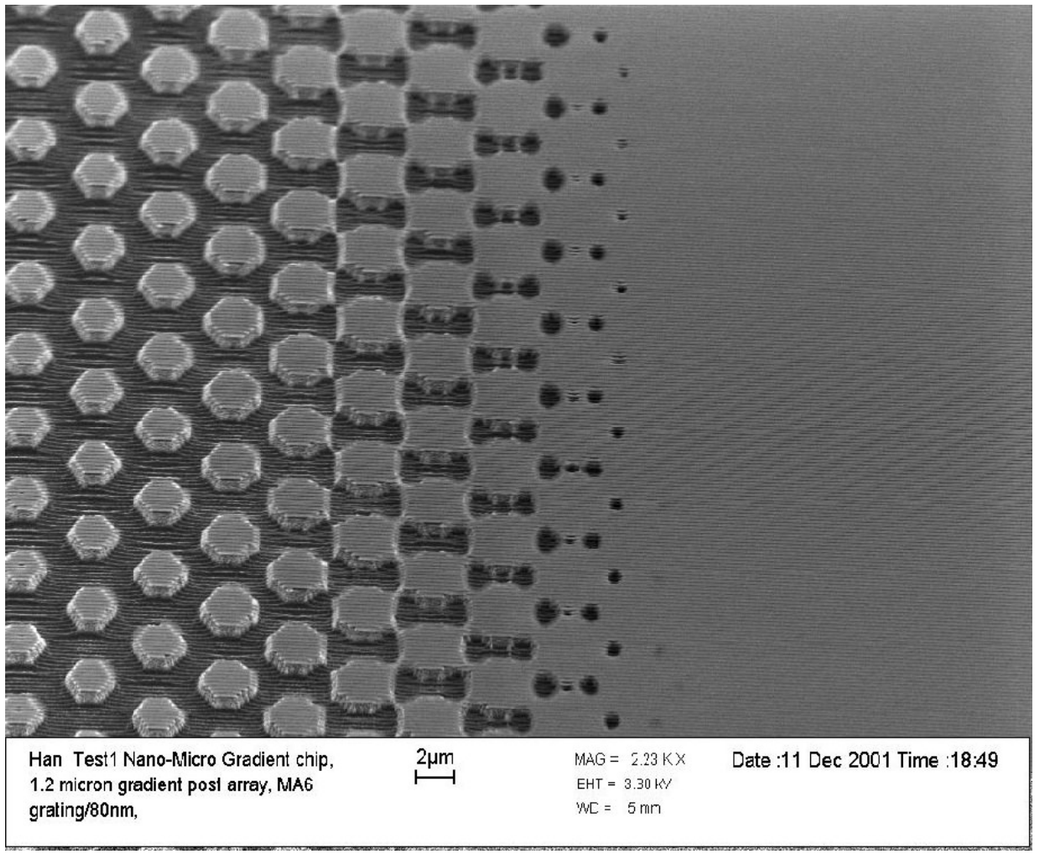




Entropy barrier

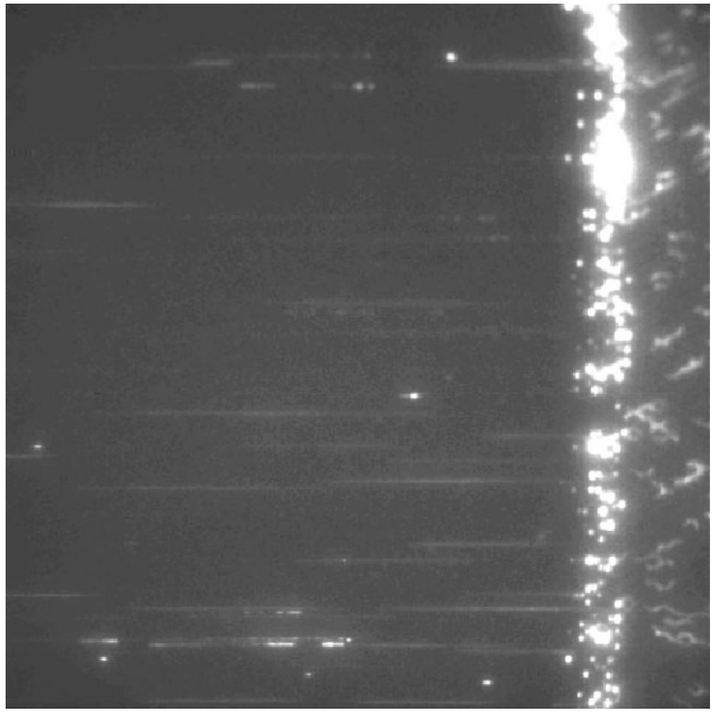
High Throughput Gradient Nanofluidic Chip



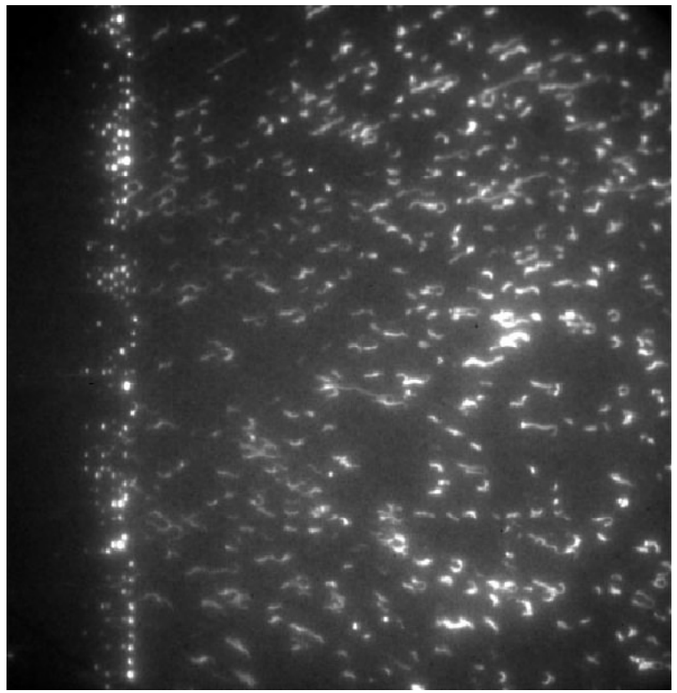


Top View (SEM)

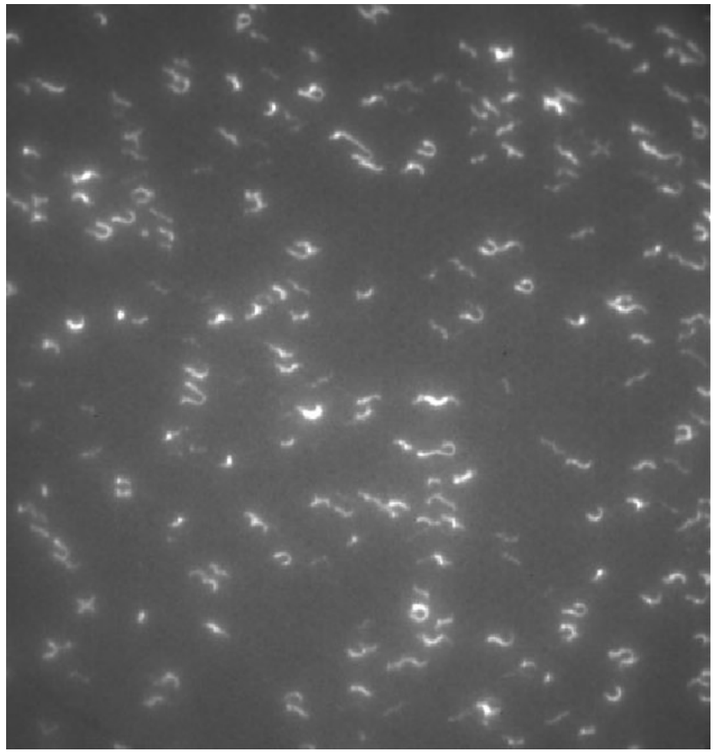
Genomic DNA Analysis On Micro/ Nano-Gradient Chip



Entering nanochannels



Approaching nanochannels



Prestreched in post array

Remember
To show Movies !

Where will the future lead us?

(1) MASSIVE amounts of detailed information about the epigenetic states of single cells, as a function of time and body state. How can we annotate all this information?

(2) Detailed network modeling of the complex interactions between protein networks in the cell. Can we have predictive power?

(3) Abilities to do genetic intervention: the insertion of genomically modified material into the genome to correct "mistakes".

(4) An end to aging: much of aging seems genetically controlled and if not reversible then at least turned off at some stage in development.

I am extremely grateful for the support I have received from the NSF over the last few years (and DARPA and the NIH too!) which allowed us to make a far more aggressive attack on my pipedreams than I thought would be possible. We are still getting better, but we are 1/3 way there.