

***RANDOM THOUGHTS ABOUT
RANDOM KNOTTING:***

***LATTICES AND TRAJECTORIES:
RAY AND STU***

De Witt Sumners

Department of Mathematics

Florida State University

Tallahassee, FL 32306

sumners@math.fsu.edu

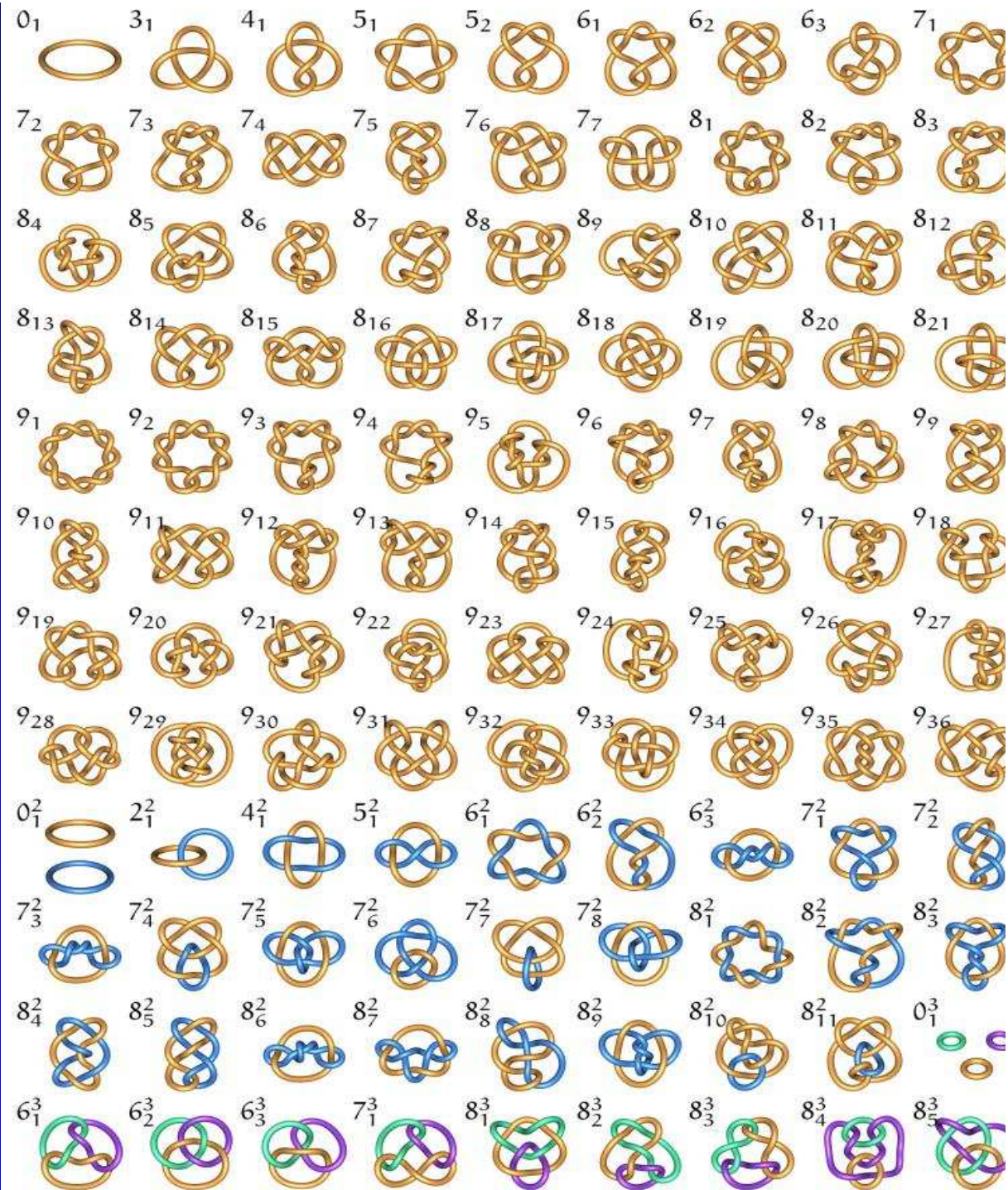
RANDOM KNOTTING

- Proof of the Frisch-Wasserman-Delbruck conjecture--the longer a random circle, the more likely it is to be knotted
- DNA knotting in viral capsids

A Knot Zoo

by Robert G. Scharein

<http://www.pims.math.ca/knotplot/zoo/>



TOPOLOGICAL ENTANGLEMENT IN POLYMERS

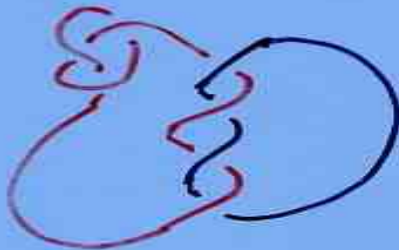
**PROBLEM: FIND SIZE-DEPENDENT
STATISTICAL MEASURES OF ENTANGLEMENT
COMPLEXITY**



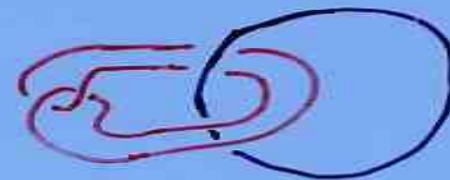
WRITHING



KNOTTING



LINKING

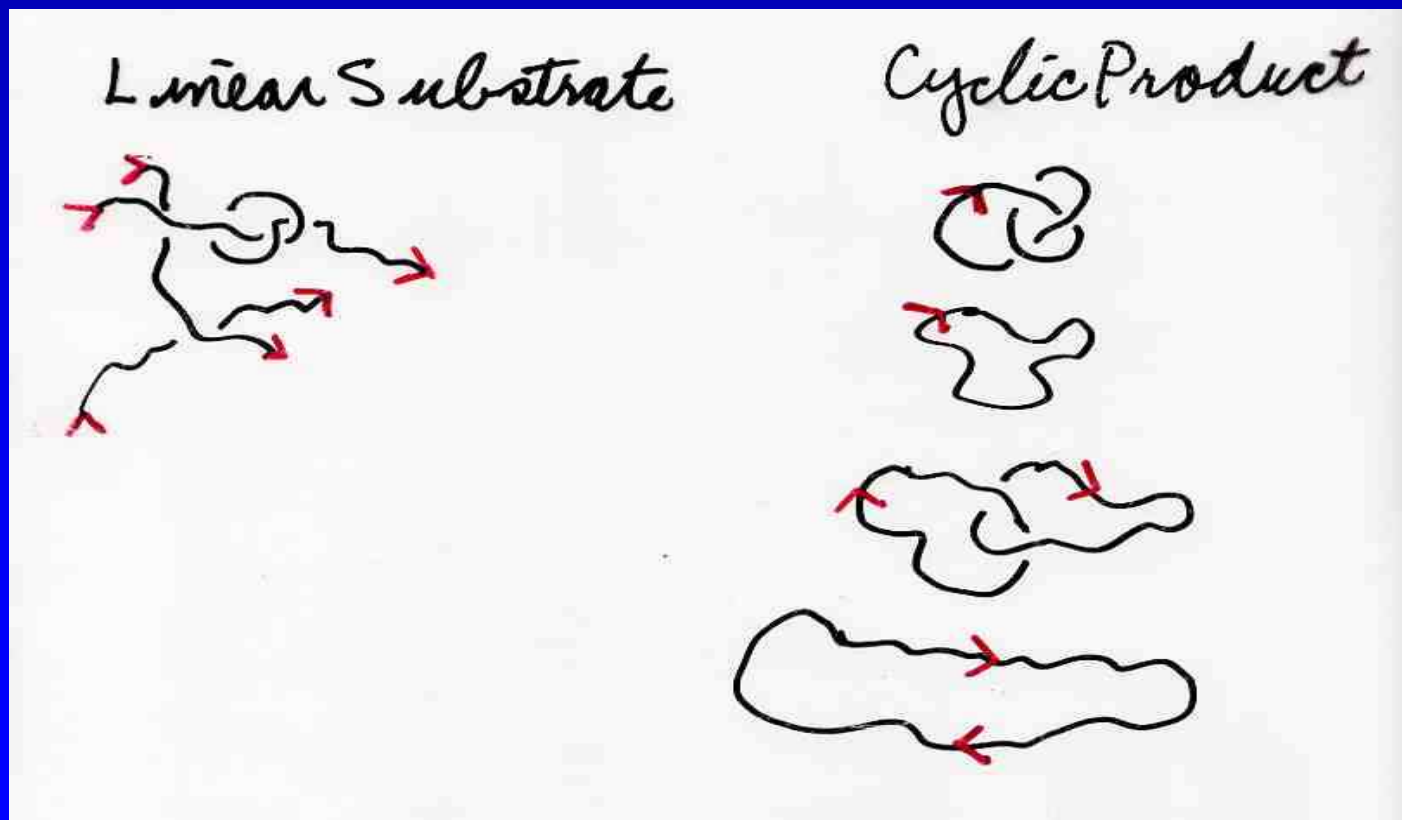


WHY STUDY RANDOM ENTANGLEMENT?

- Polymer chemistry and physics: microscopic entanglement related to macroscopic chemical and physical characteristics--flow of polymer fluid, stress-strain curve, phase changes (gel formation)
- Biopolymers: entanglement encodes information about biological processes--random entanglement is experimental noise and needs to be subtracted out to get a signal

BIOCHEMICAL MOTIVATION

Predict the yield from a random cyclization experiment in a dilute solution of linear polymers



MATHEMATICAL PROBLEM

- If L is the length of linear polymers in dilute solution, what is the yield (the spectrum of topological products) from a random cyclization reaction?
- L is the # of repeating units in the chain--# of monomers, or # of Kuhn lengths (equivalent statistical lengths)--for polyethylene, Kuhn length is about 3.5 monomers. For duplex DNA, Kuhn length is about 300-500 base pairs

FRISCH-WASSERMAN- DELBRUCK CONJECTURE

- L = # edges in random polygon
- $P(L)$ = knot probability

$$\lim_{L \rightarrow \infty} P(L) = 1$$

Frisch & Wasserman, JACS 83(1961), 3789
Delbruck, Proc. Symp. Appl. Math. 14 (1962), 55

ENTER STU WHITTINGTON!

- CIC Symposium on *Computational and Mathematical Chemistry*, Saskatoon, June 1986.
- Nail: FWD Conjecture
- Hammer: Kesten Pattern Theorem

PROOF OF FWD CONJECTURE

THEOREM:

$$P(L) \sim 1 - \exp(-\lambda L) \quad \lambda > 0$$

Sumners & Whittington, J. Phys. A: Math. Gen. 23
(1988), 1689

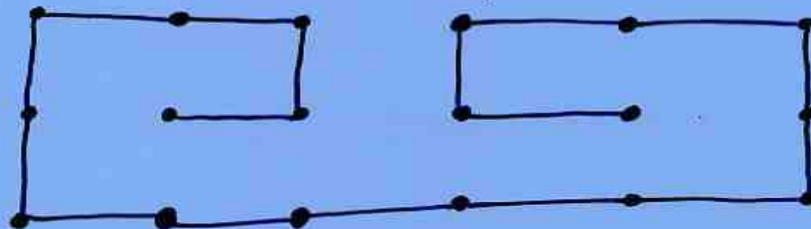
Pippenger, Disc Appl. Math. 25 (1989), 273

KESTEN PATTERNS

KESTEN PATTERN:



NOT A KESTEN PATTERN:



Kesten, J. Math. Phys. 4(1963), 960

TIGHT KNOTS

THEOREM:

*T ANY KESTEN PATTERN; $\exists \alpha > 0$ SUCH THAT
T APPEARS AT LEAST αL TIMES IN ALL BUT
EXPONENTIALLY FEW SAW (SAP) OF
LENGTH L .*

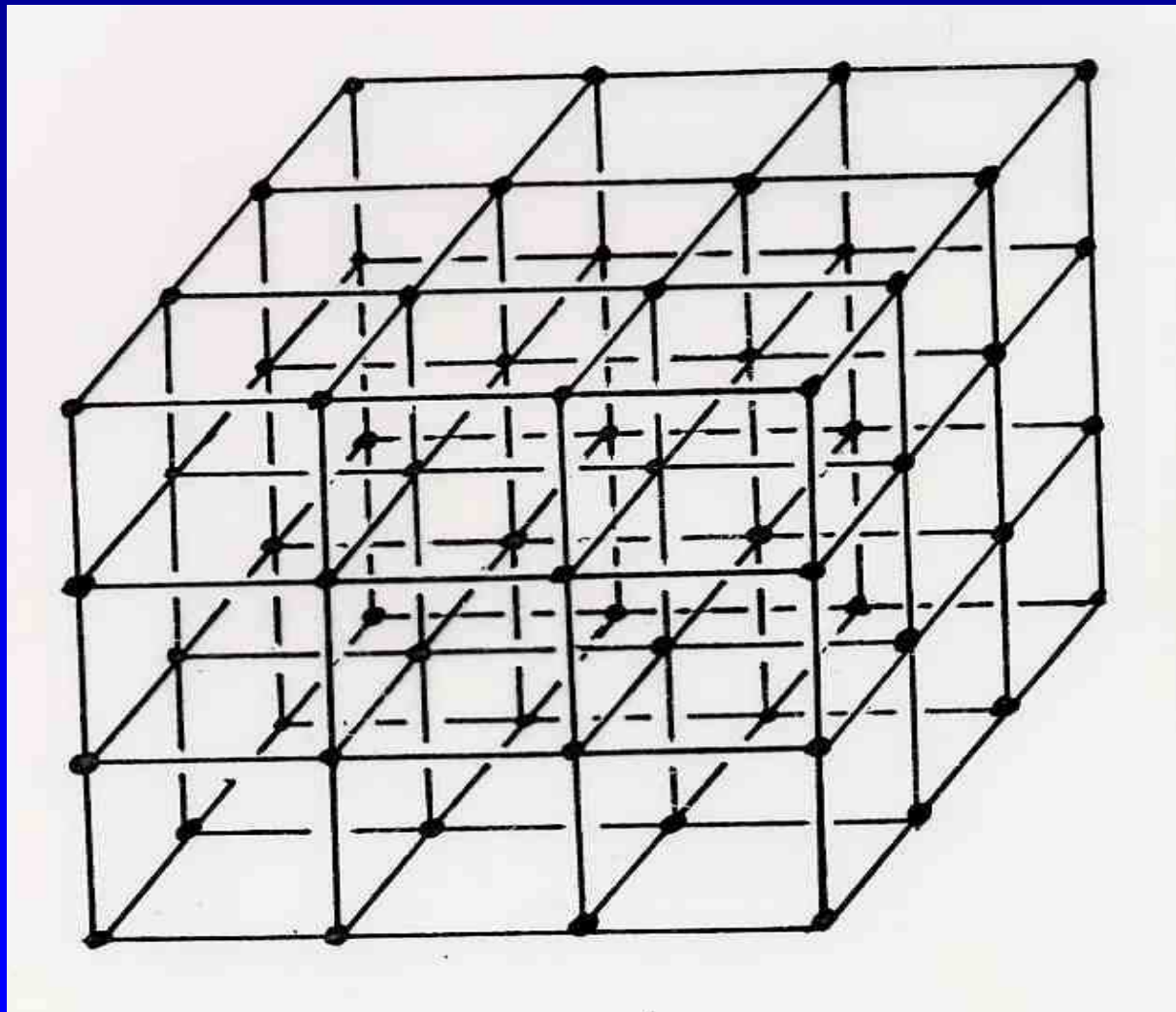
IDEA OF PROOF:

**PRODUCE A KESTEN PATTERN T SUCH THAT
IF T APPEARS AT LEAST ONCE IN A SAW,
THEN THE SAW IS KNOTTED.**

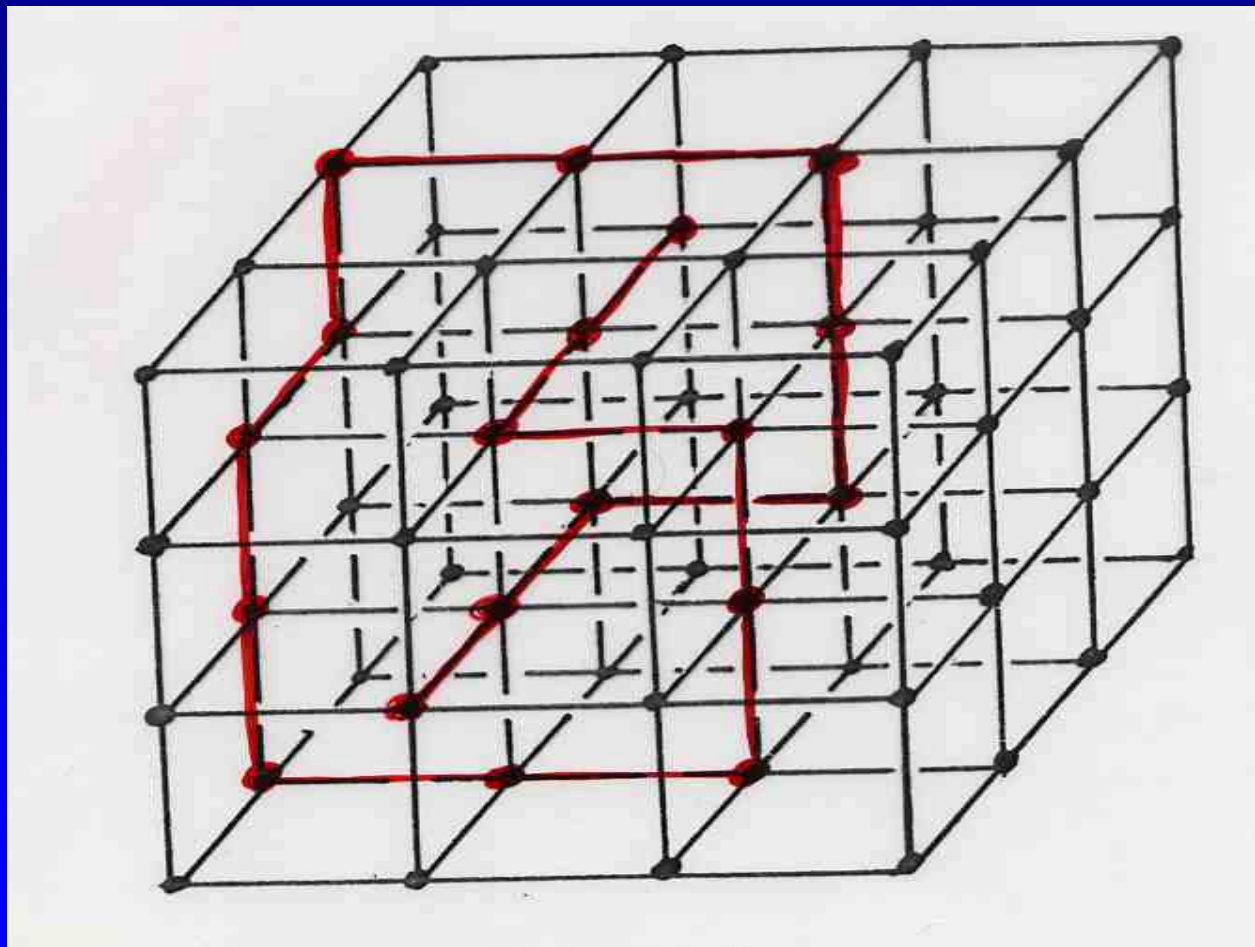
T = TIGHT KNOT !



\mathbb{Z}^3



TIGHT KNOT ON Z^3

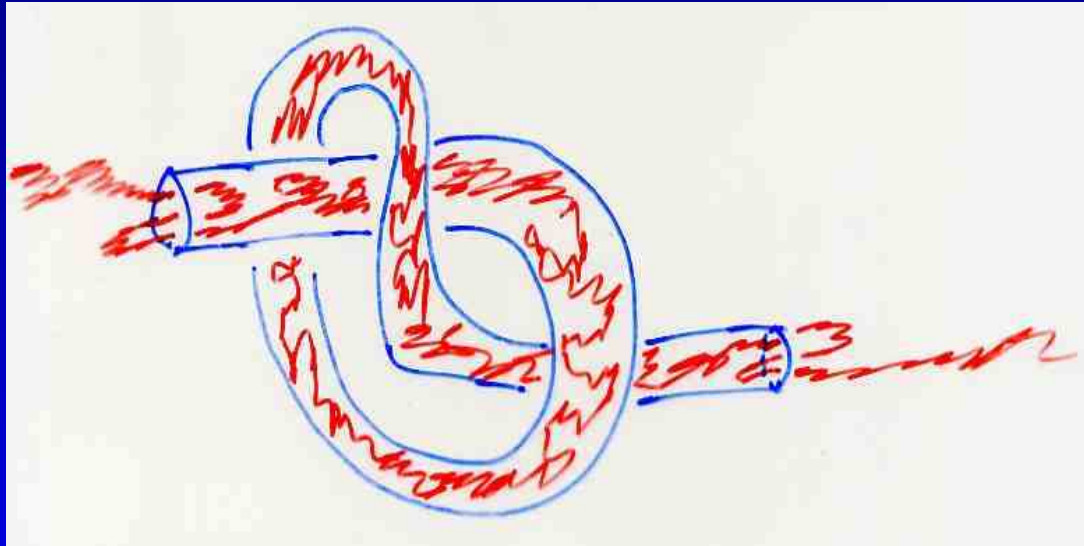


19 vertices, 18 edges

RANDOM KNOT QUESTIONS

- For fixed length n , what is the distribution of knot types?
- How does this distribution change with n ?
- What is the asymptotic behavior of knot complexity--growth laws $\sim \beta n^\alpha$?
- How to quantize entanglement of random arcs?

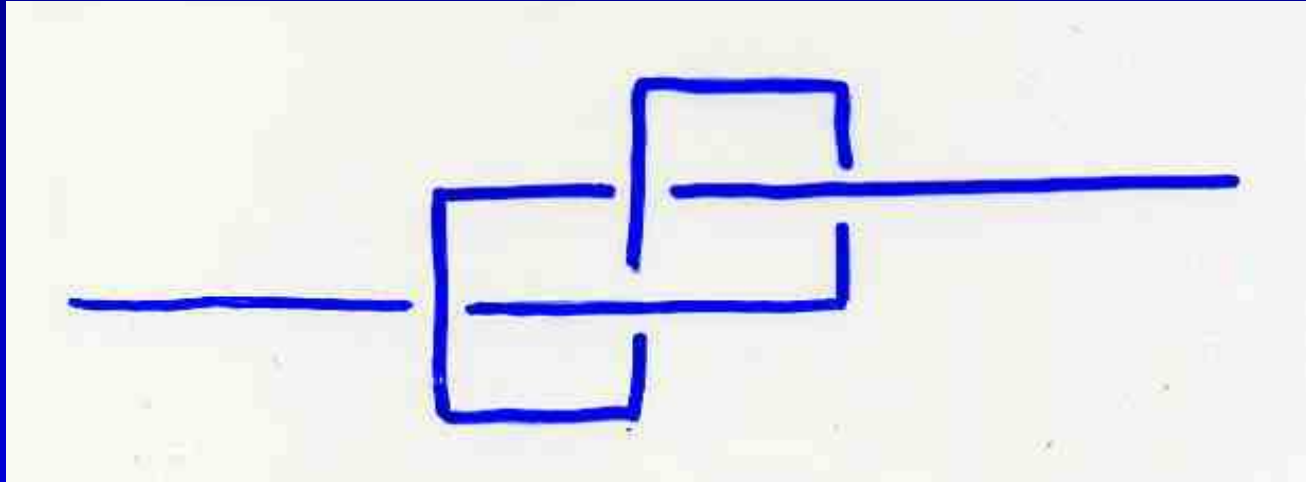
KNOTS IN BROWNIAN FLIGHT



- *All knots at all scales*

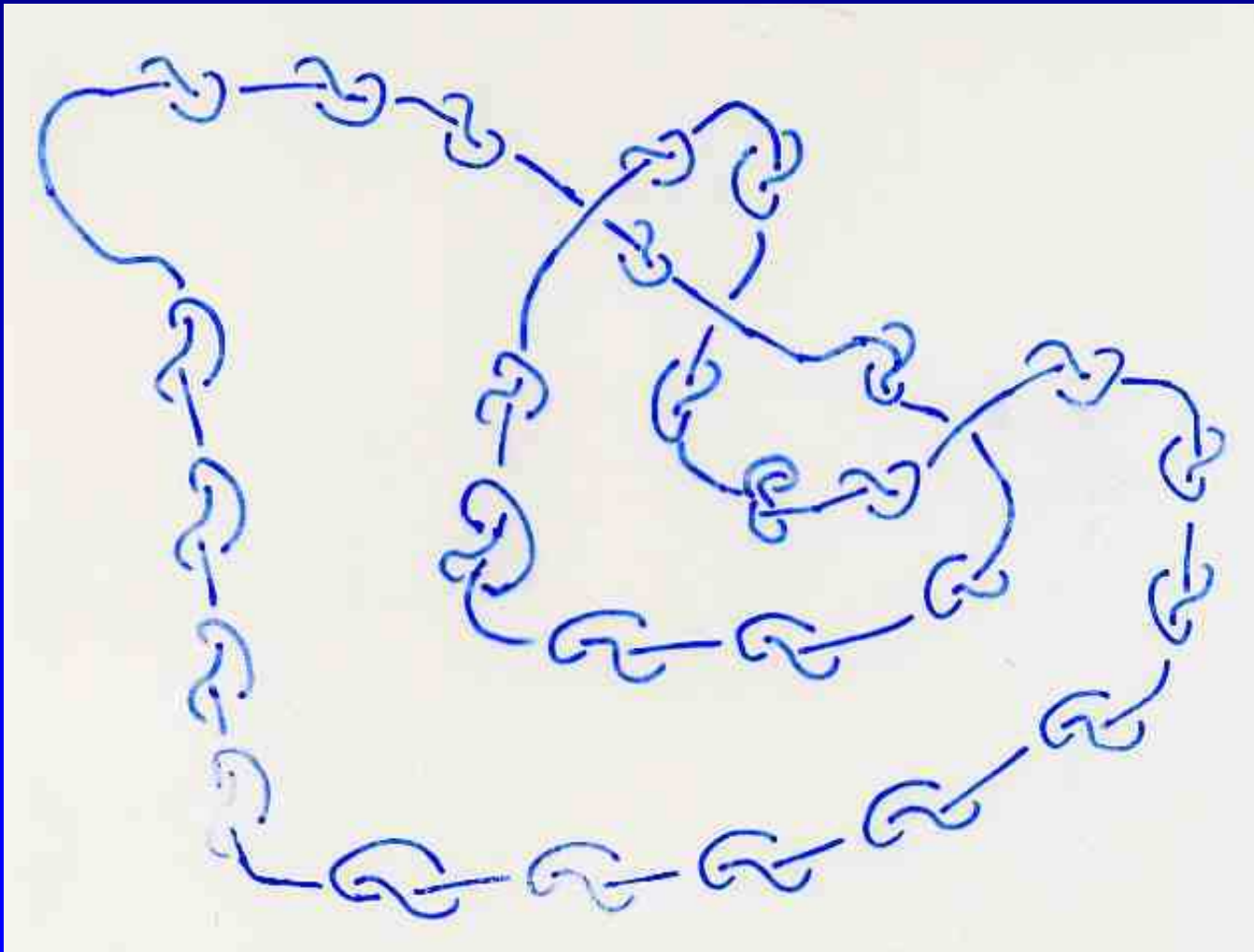
Kendall, J. Lon. Math. Soc. 19 (1979), 378

ALL KNOTS APPEAR



Every knot type has a tight Kesten pattern representative on \mathbb{Z}^3

LONG RANDOM KNOTS



MEASURING KNOT COMPLEXITY

\mathcal{K} = THE SET OF KNOT TYPES

$F: \mathcal{K} \rightarrow [0, \infty)$ IS A

GOOD MEASURE OF KNOT COMPLEXITY

IF

(i) $F(\text{UNKNOT}) = 0$

(ii) $\exists K \in \mathcal{K}$ SUCH THAT

$$F(nK \# L) \geq nF(K) > 0$$

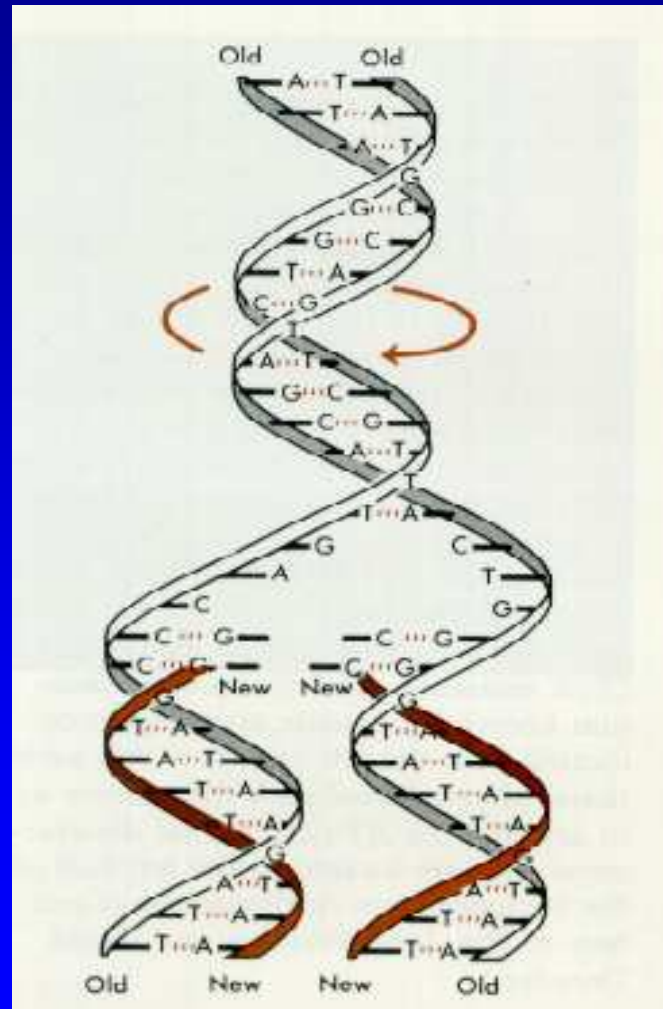
for all $L \in \mathcal{K}$.

LONG RANDOM KNOTS ARE VERY COMPLEX

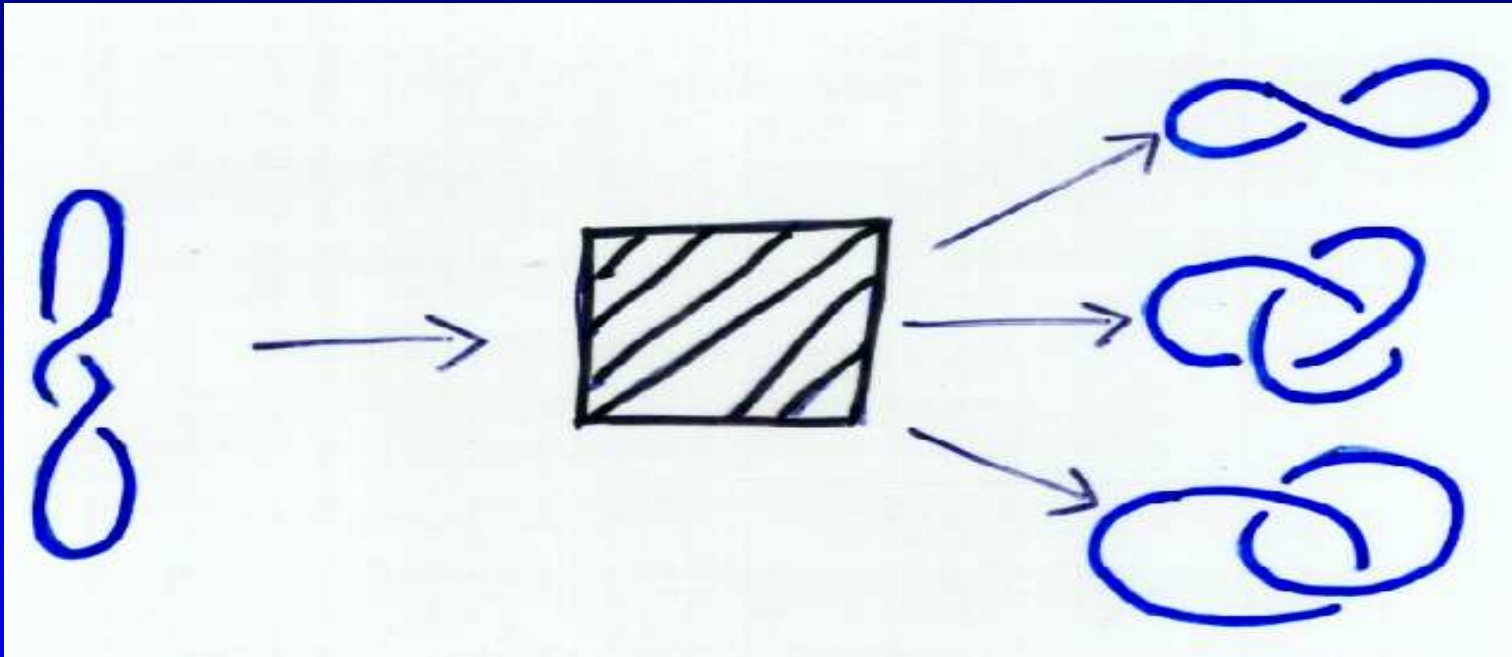
THEOREM: All good measures of knot complexity diverge to $+\infty$ at least linearly with the length--the longer the random polygon, the more entangled it is.

Examples of good measures of knot complexity:
crossover number, unknotting number, genus, bridge number, braid number, span of your favorite knot polynomial, total curvature, etc.

DNA Replication



Topological Enzymology

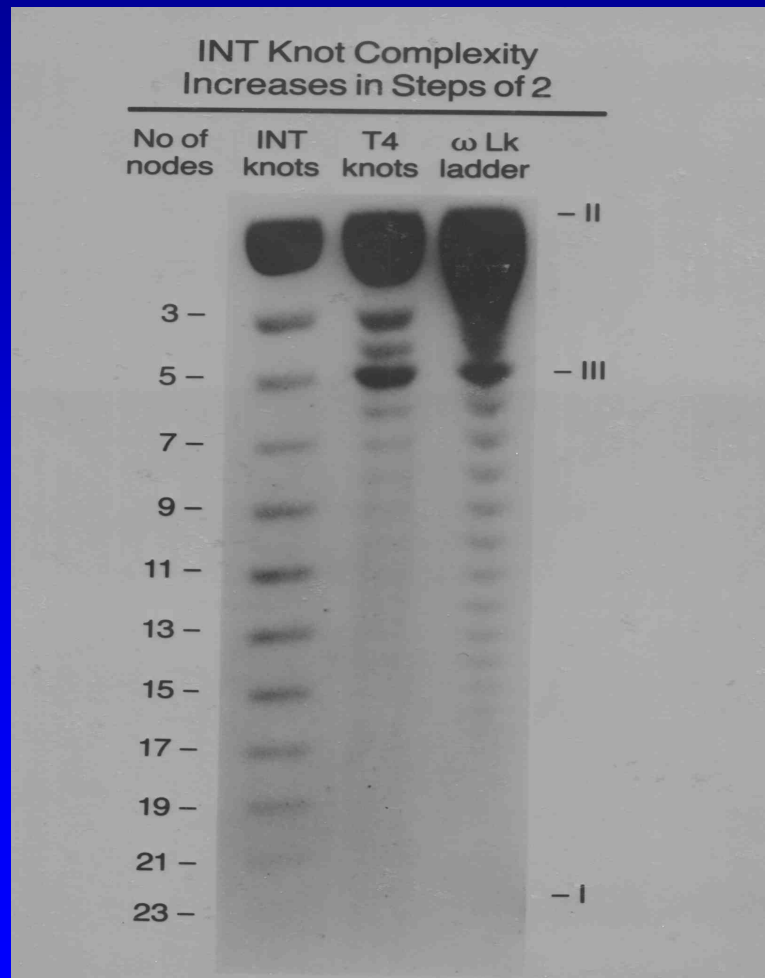


Mathematics: Deduce enzyme binding and mechanism from observed products

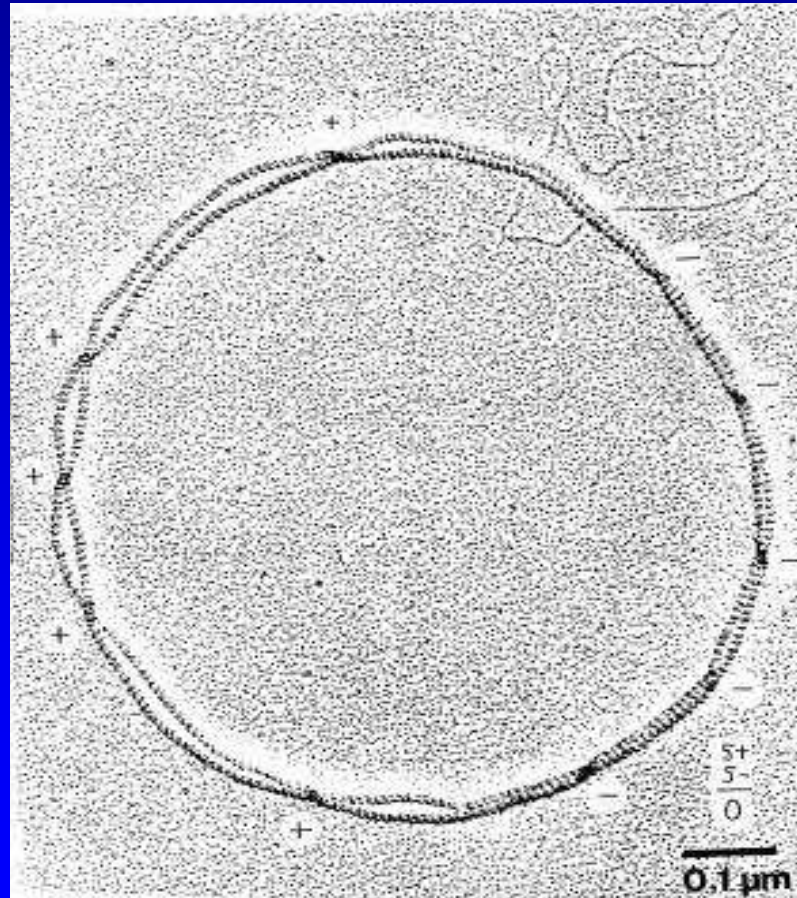
TOPOLOGICAL ENZYMOLOGY

- React circular DNA plasmids *in vitro* (*in vivo*) with purified enzyme
- Gel electrophoresis to separate products (DNA knots & links)
- Electron microscopy of RecA coated products
- Use topology and geometry to build predictive models

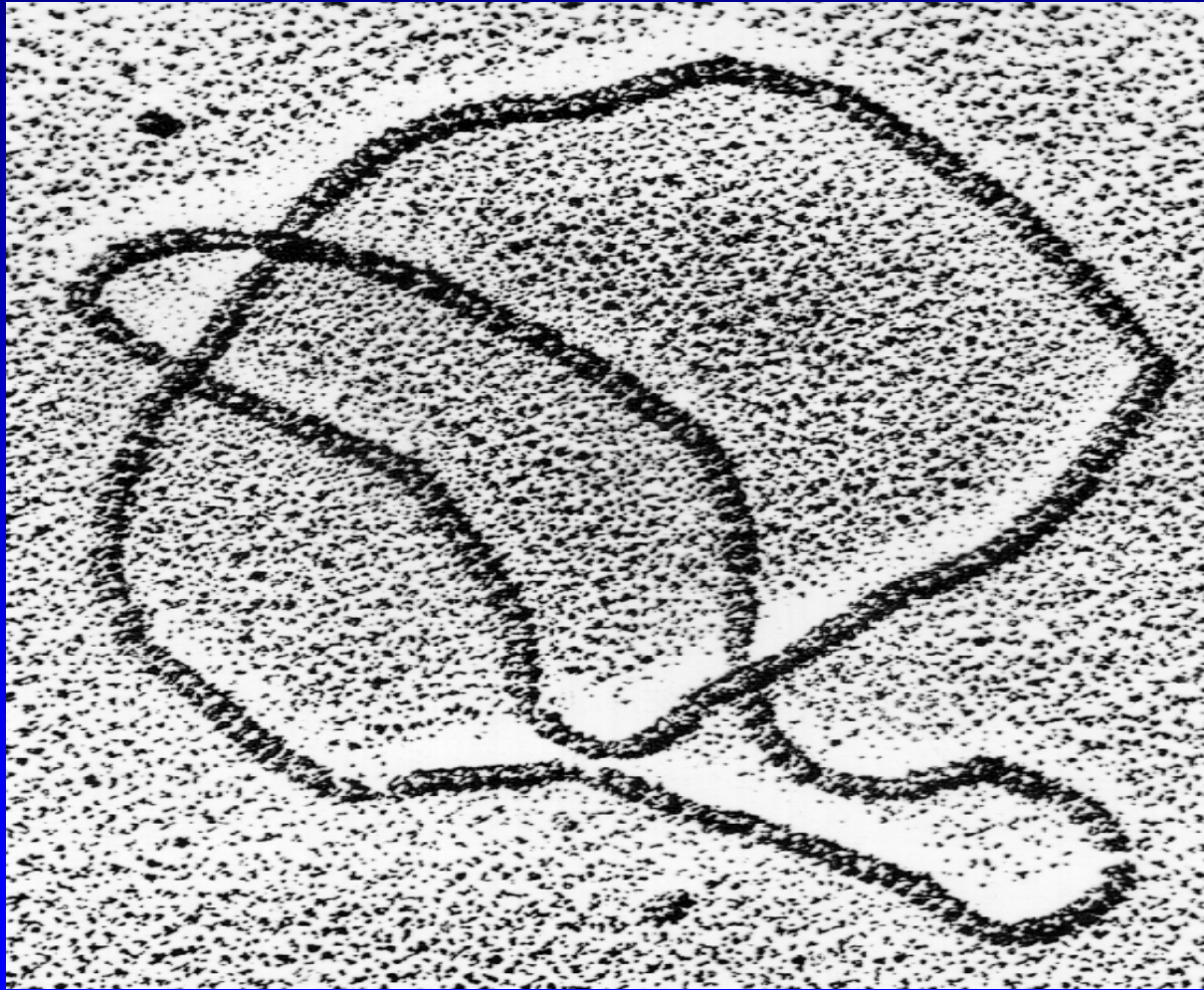
GEL ELECTROPHORESIS



RecA Coated DNA

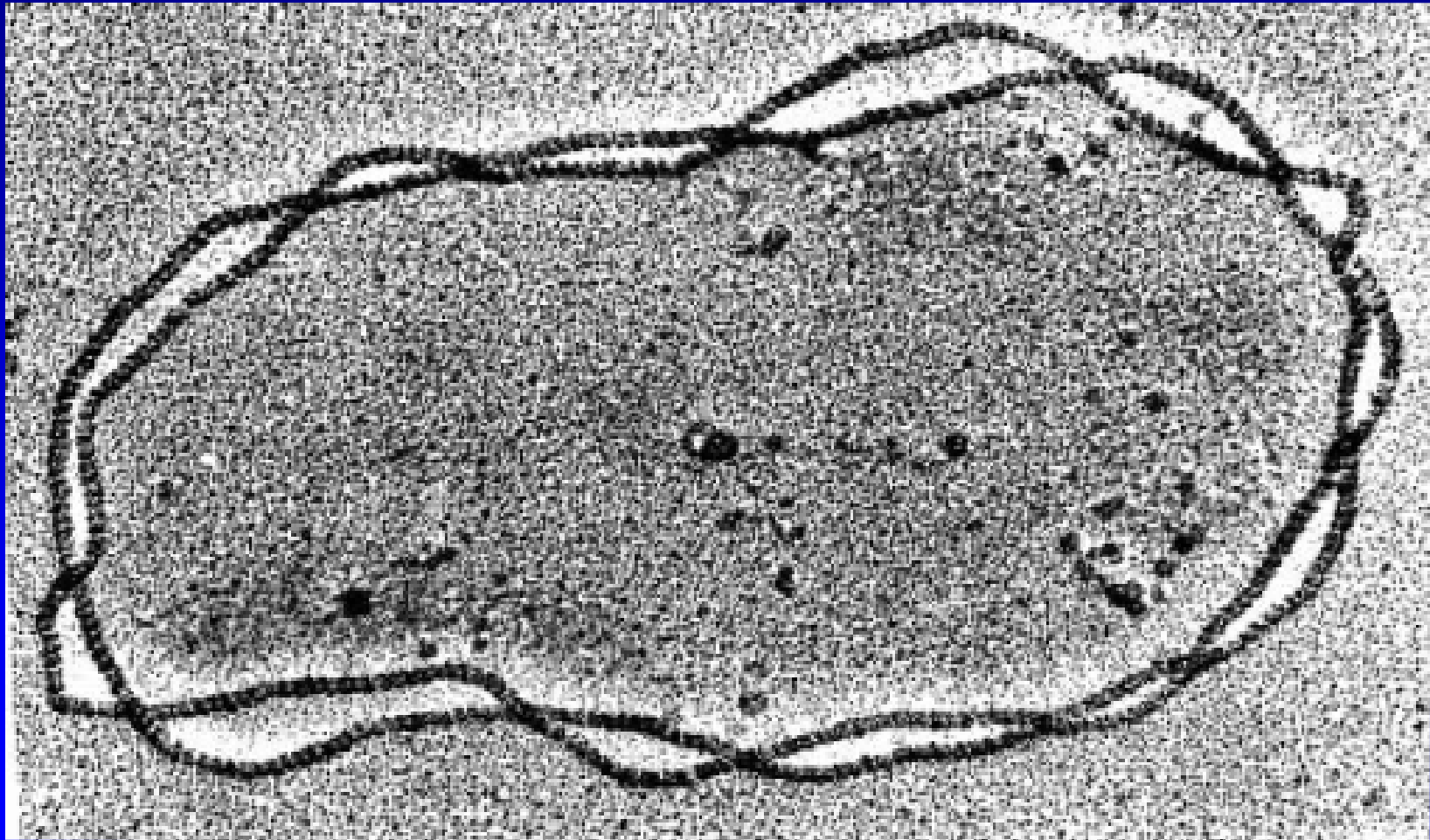


DNA Trefoil Knot



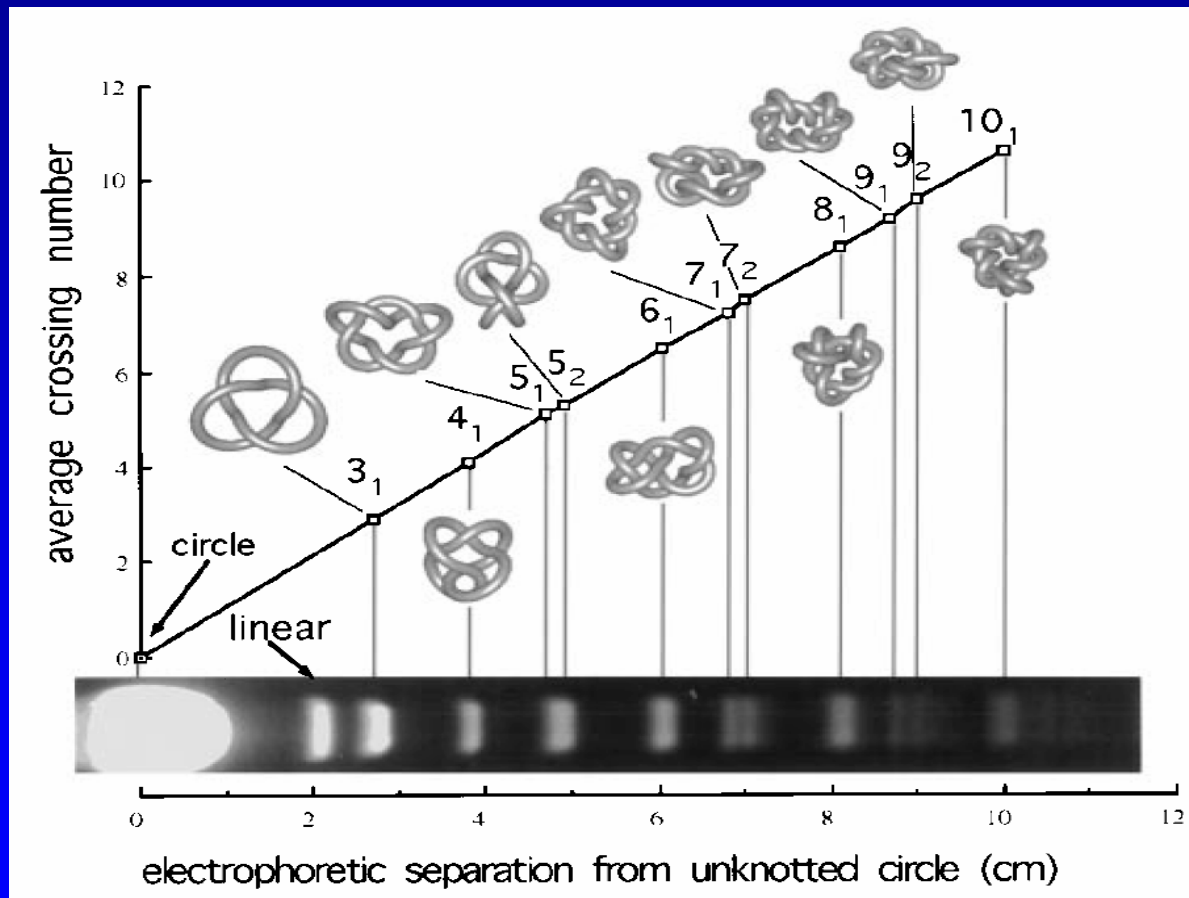
Dean et al., J BIOL. CHEM. 260(1985), 4975

DNA (2,13) TORUS KNOT



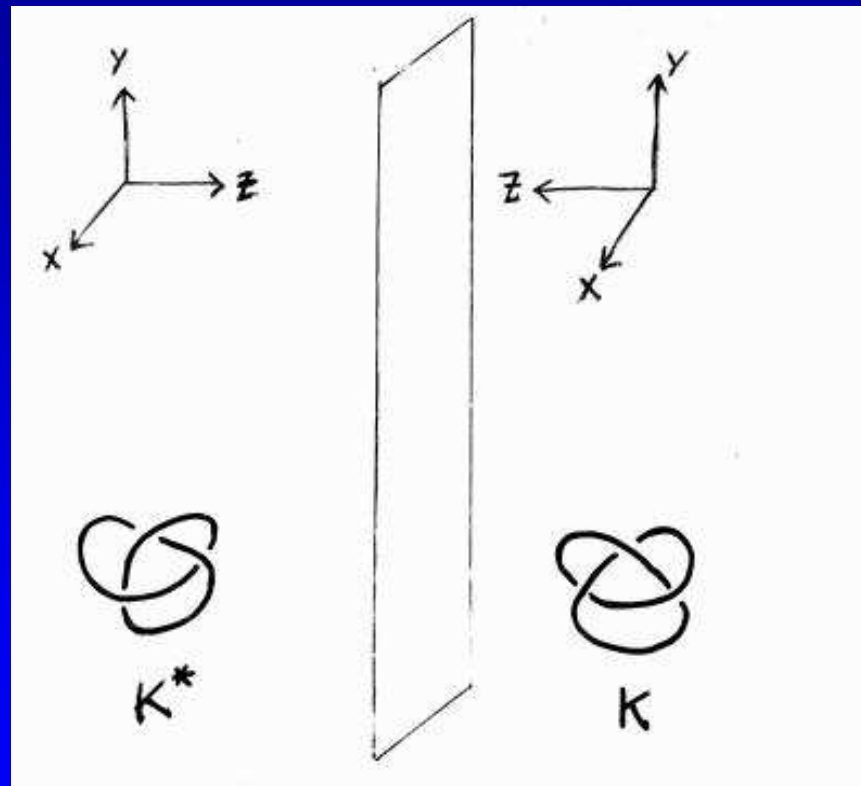
Spengler et al. CELL 42 (1985), 325

GEL VELOCITY IDENTIFIES KNOT COMPLEXITY

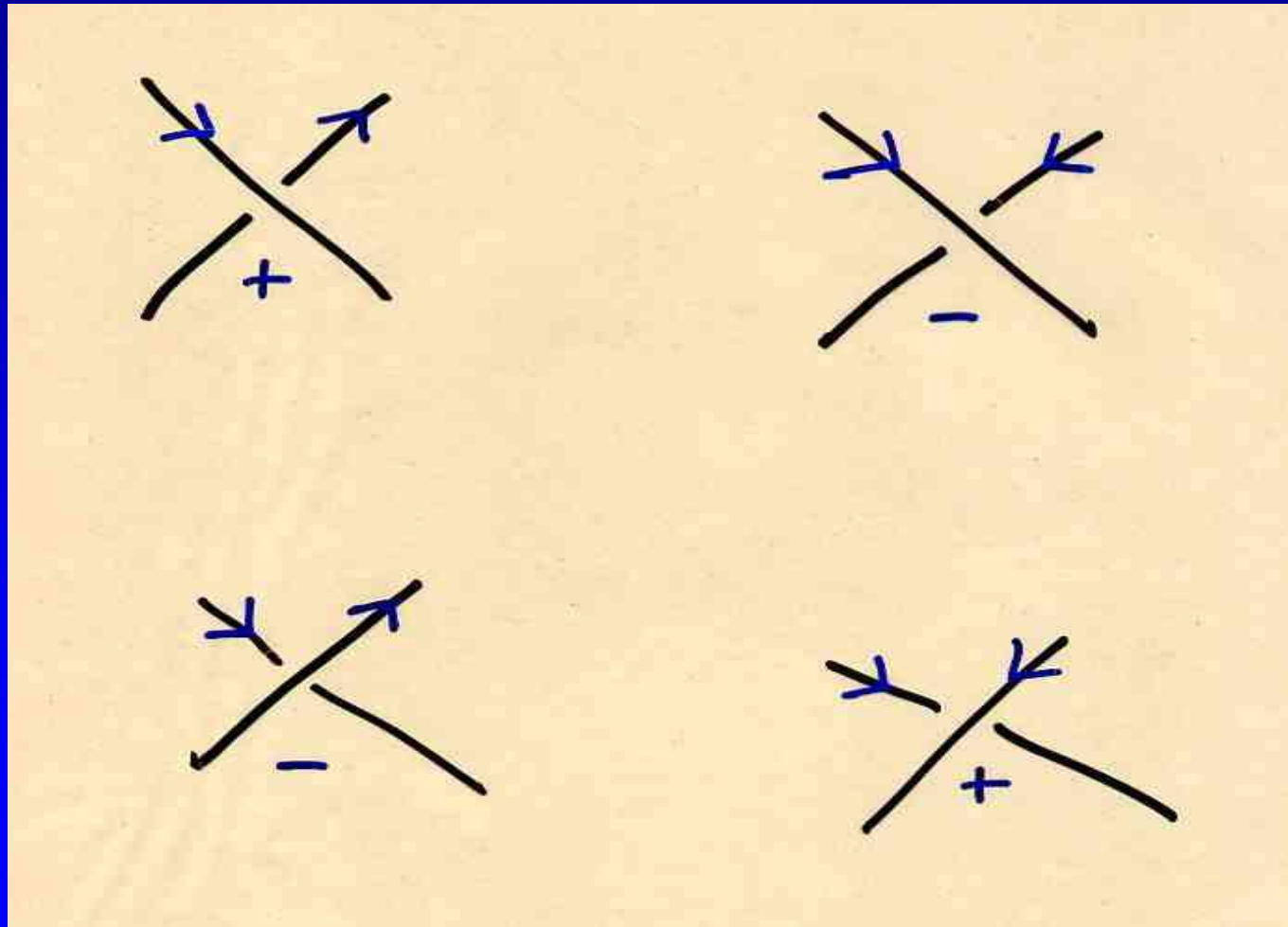


Vologodskii et al, JMB 278 (1988), 1

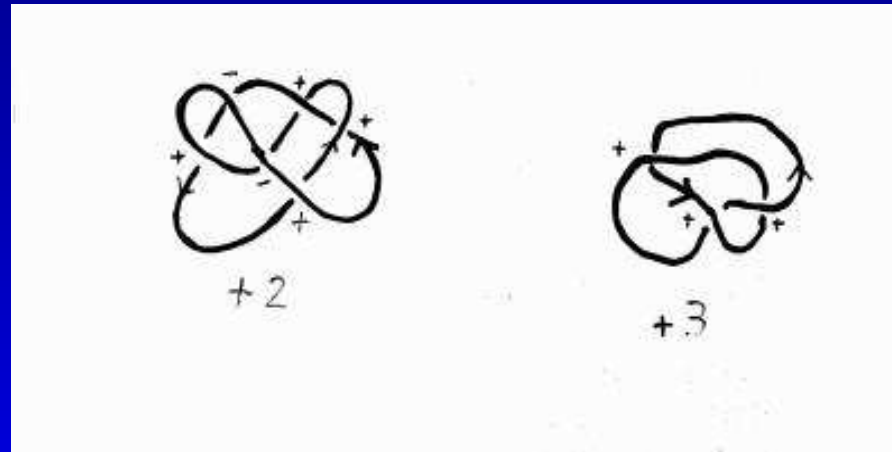
CHIRALITY



CROSSING SIGN CONVENTION

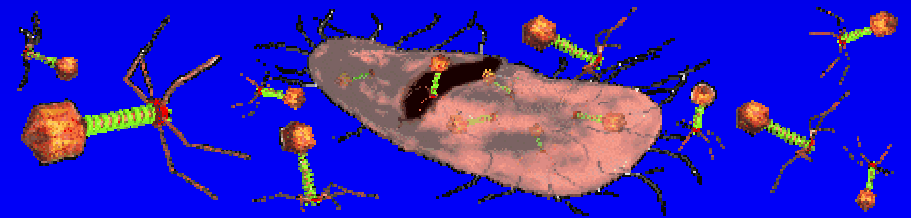
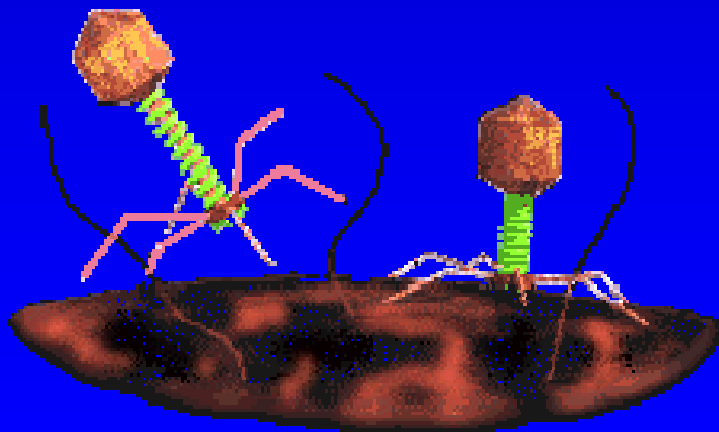
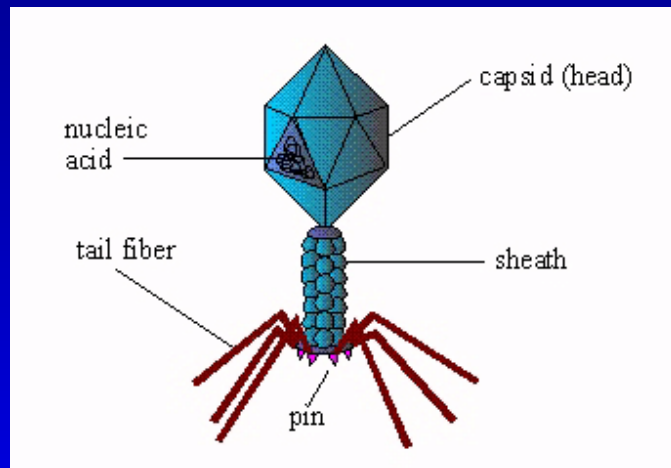


WRITHE & AVERAGE CROSSING NUMBER



Writhe --average the sum of signed crossings over all projections (average number of crossings over all projections)

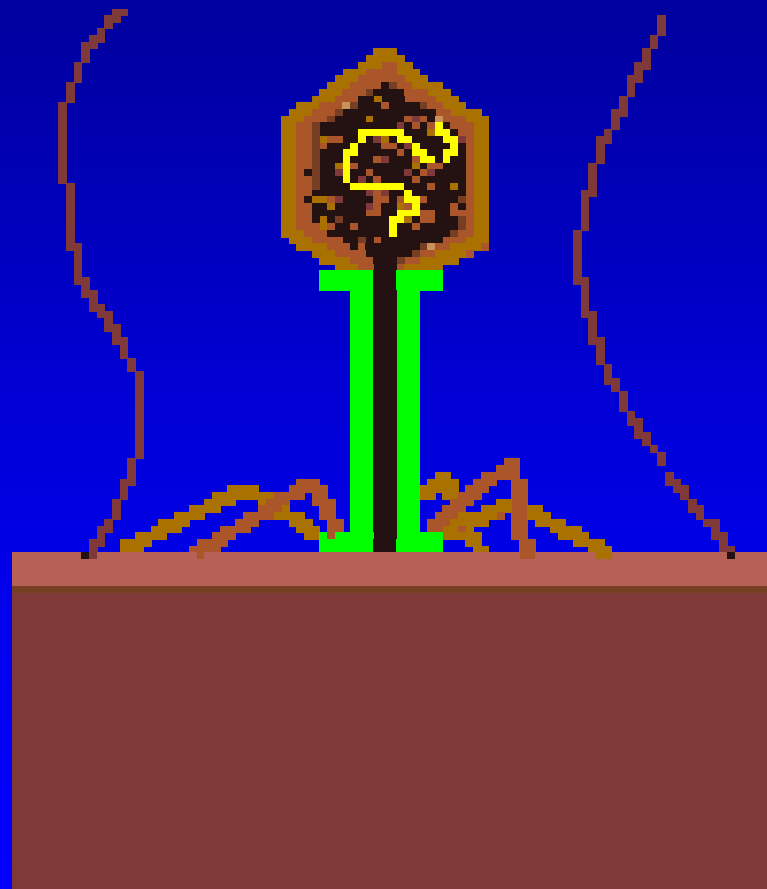
VIRUS LIFE CYCLE



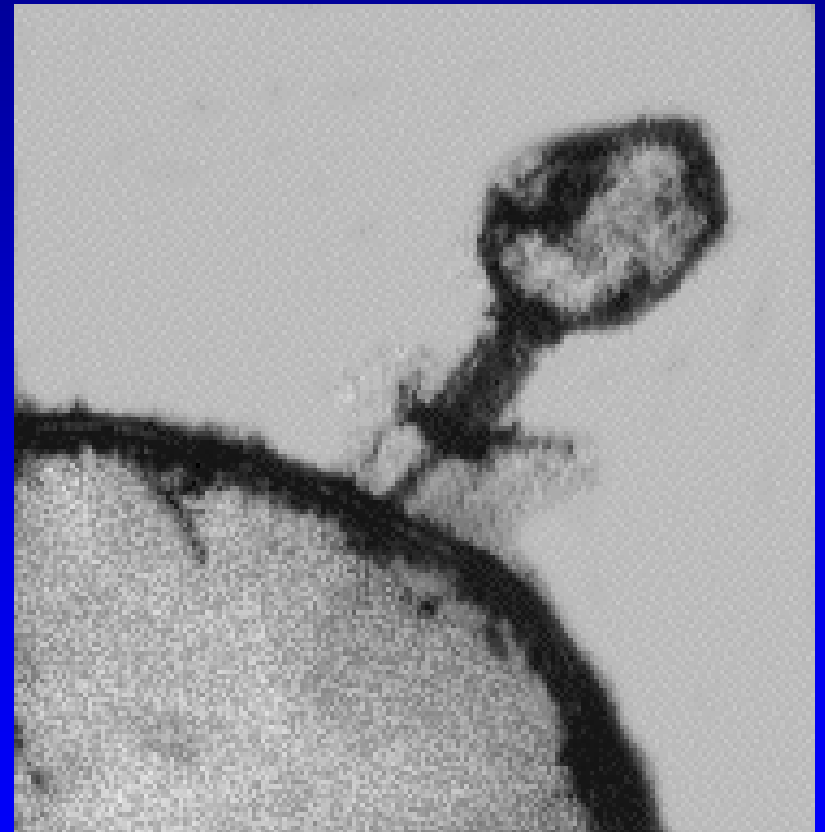
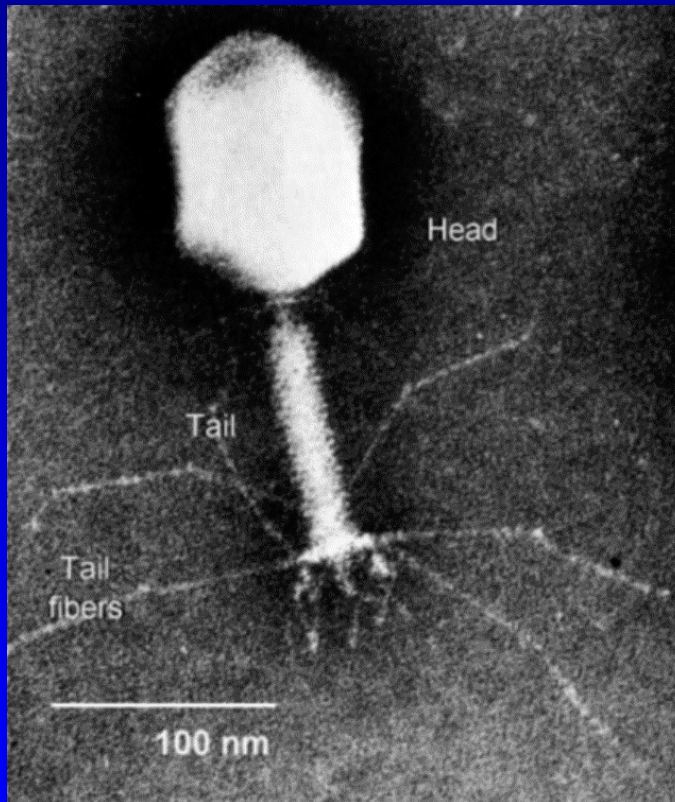
VIRUS ATTACKS!



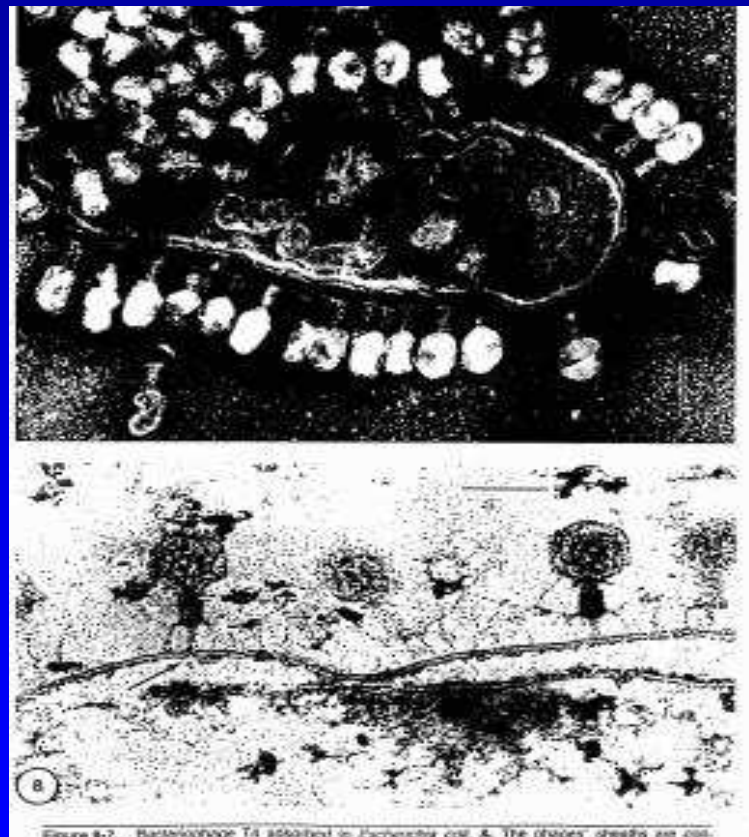
VIRUS ATTACKS!



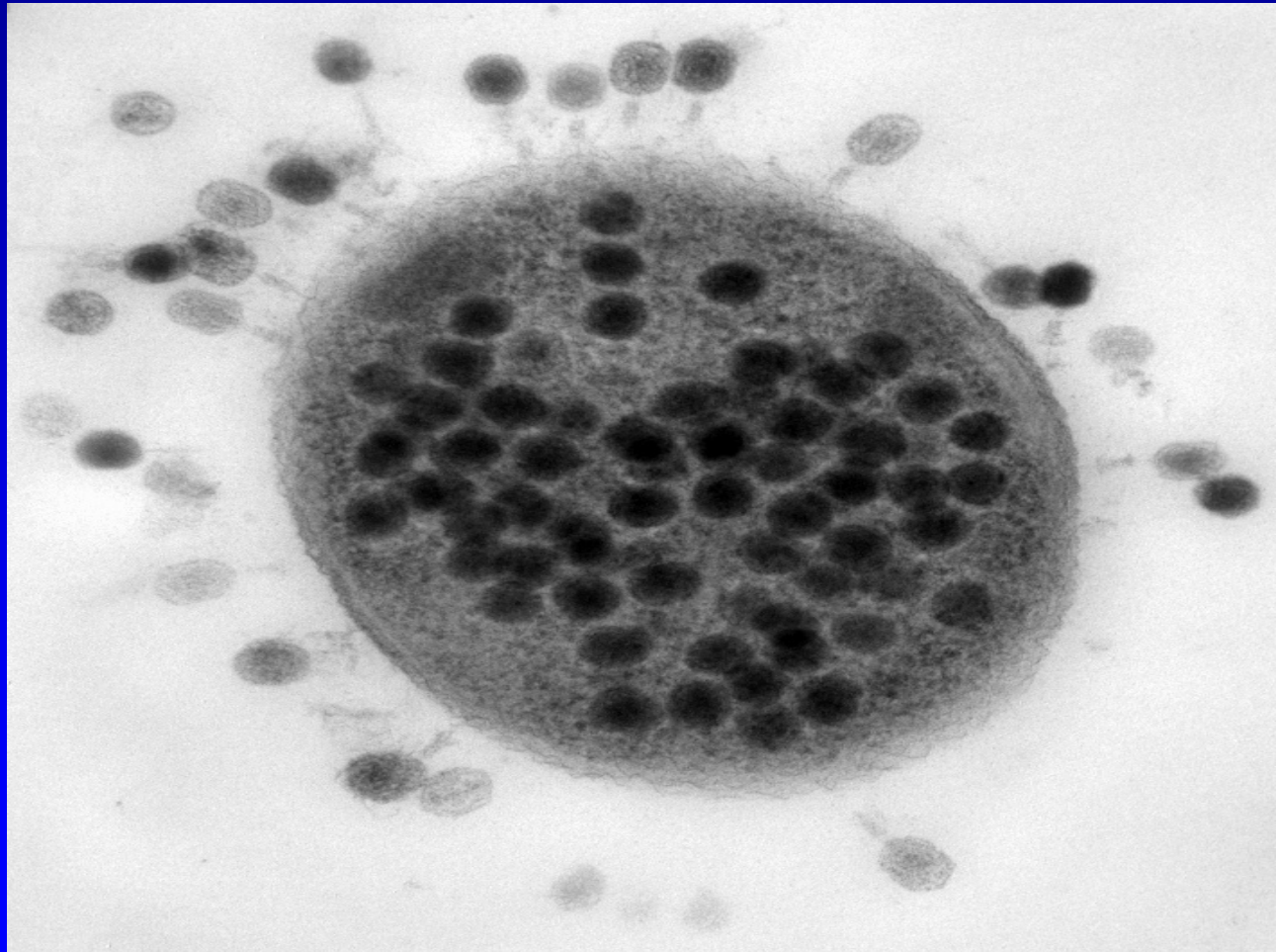
T4 EM



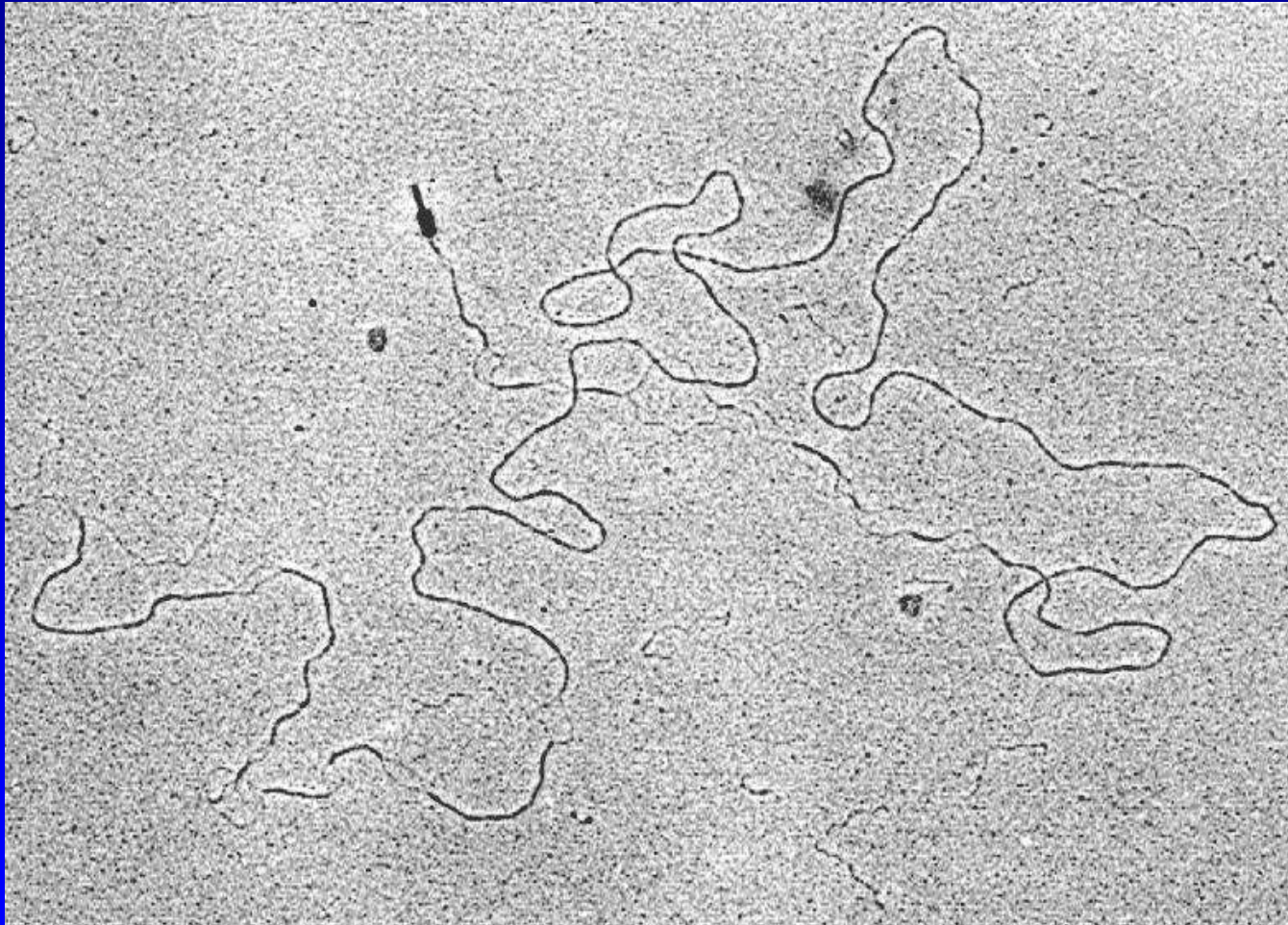
VIRUS ATTACK



T4 ATTACK



HOW IS THE DNA PACKED?



SPOOLING MODEL

Model proposed by :

Richards et al. J. Mol. Biol. 78 (1973) 255-259

Earnshaw et al. Nature 268 (1977) 598-602

Sewer J. Mol. Biol. 190 (1986) 509-512

Cerritelli et al. Cell 91 (1997) 271-280

Properties:

DNA is a locally parallel array

Uniform structure, continuously wound around an axis

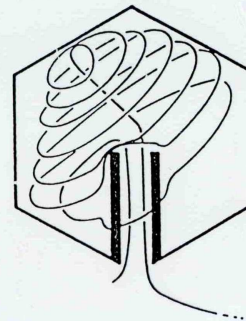
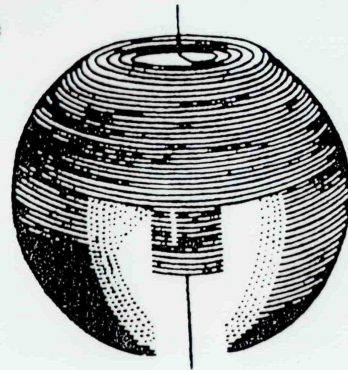
Interhelix distance 29\AA

B form

Viruses:

P22, Lambda, T7

Structure



FOLD MODEL

Model proposed by:

Richards et al. J. Mol. Biol. 78 (1973) 255-259

Black et al. Proc. Natl. Acad. Sci. USA 82 (1985) 7960-7964

Properties:

DNA is a locally parallel array

First DNA packed is in the center of the capsid

Interhelix distance in local domains 25 Å

Regions ordered parallel to the capsid axis 160-300 bp long

Kinks 3 bp long

B form

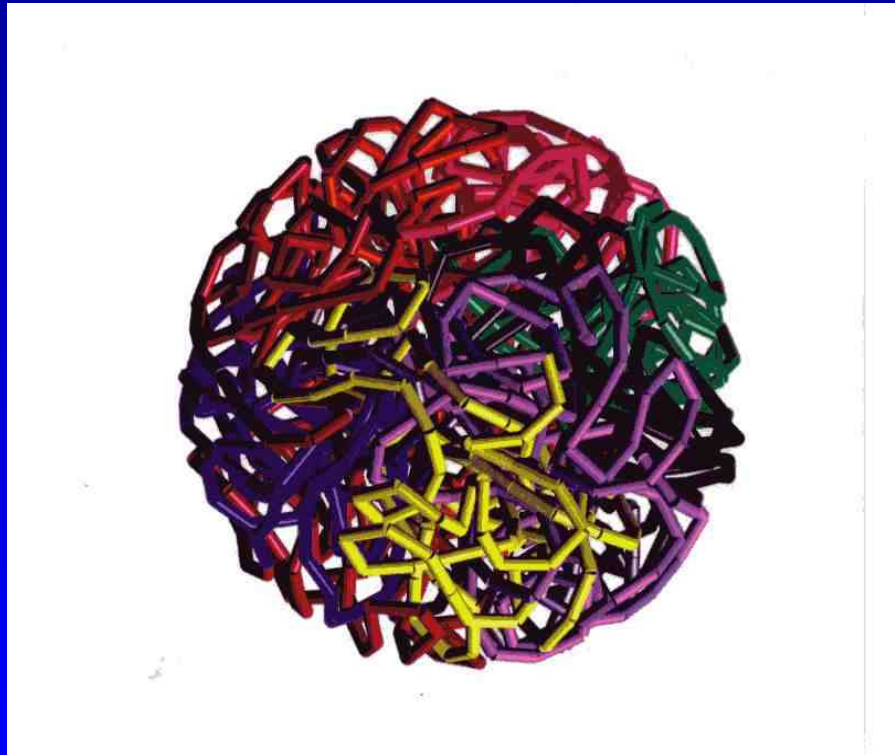
Viruses:

T4

Structure



RANDOM PACKING

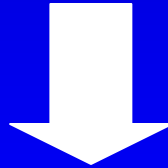


P4 DNA has cohesive ends that form closed circular molecules

GGCGAGGCGGGAAAGCAC

....

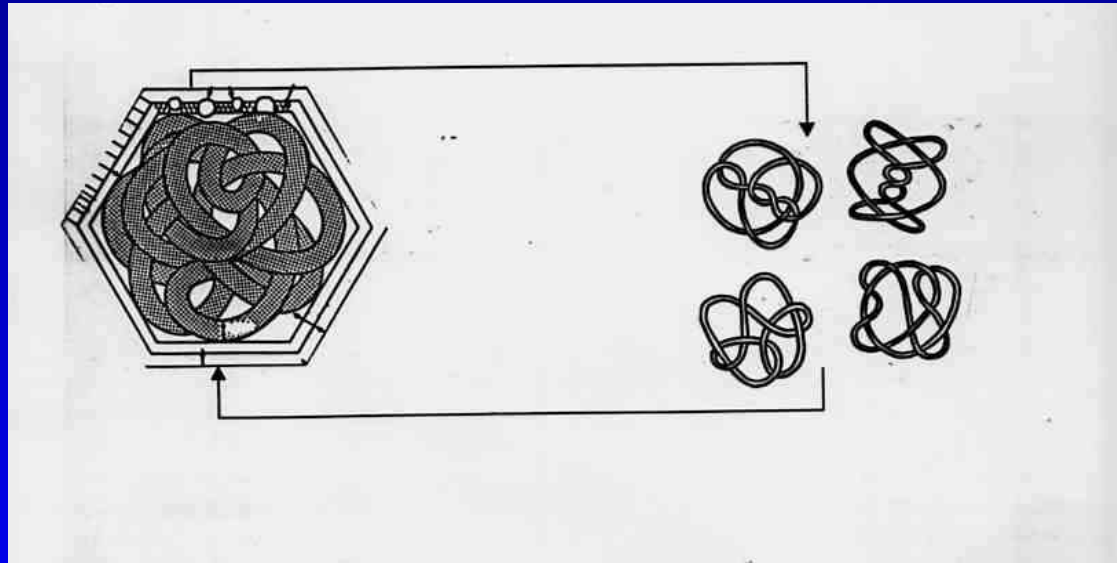
..... CCGCTCCGCCCTTTCGTG



GGCGAGGCGGGAAAGCAC

CCGCTCCGCCCTTTCGTG

VIRAL KNOTS REVEAL PACKING



- Compare observed DNA knot spectrum to simulation of knots in confined volumes

Experimental Data: Tailless vs Mature Phage Knot Probability

Liu et al. Nucleic Acids Research 9, 16 (1981)
Liu et al. Proc. Natl. Acad. Sci. USA 78 (1981)

When extracting DNA molecules from mature bacteriophages P2 and P4 linear, circular not knotted and knotted structures are found.

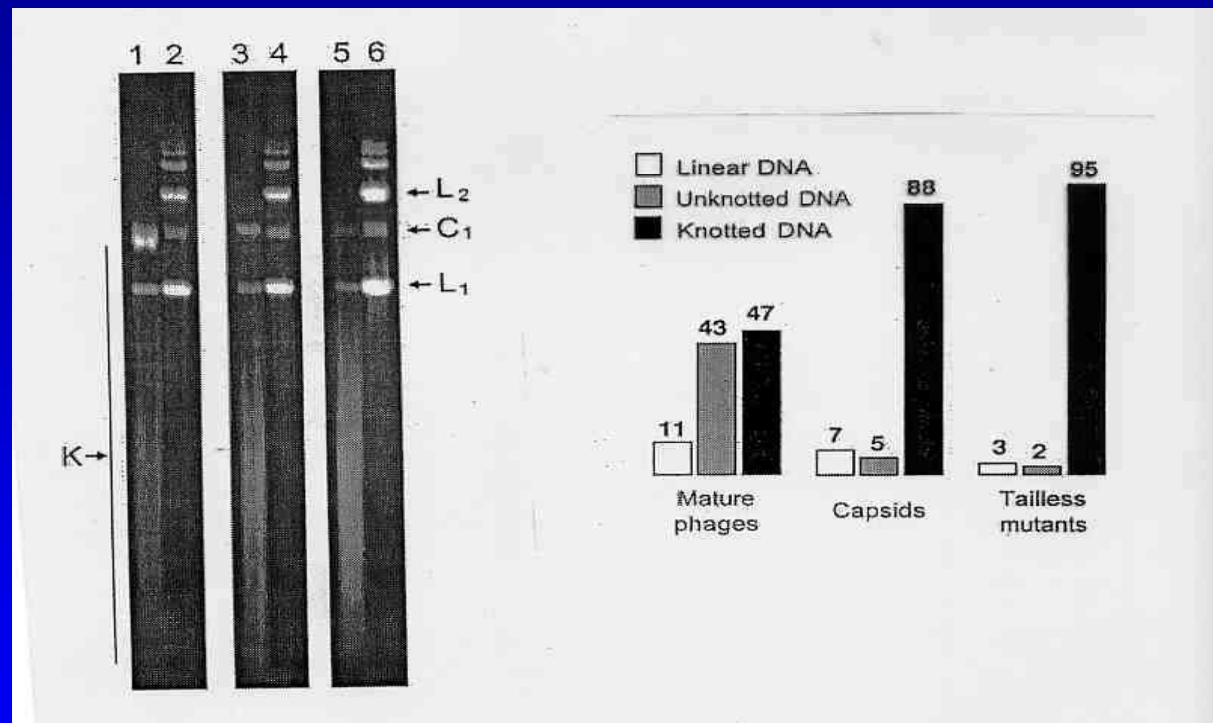
	<u>DNA length</u>	<u>Linear</u>	<u>Unknotted</u>	<u>Knotted</u>
P2vir1	33Kb	-	-	0%
P2vir22lts37	31.5Kp	11%	20%	69%
P2vir1am12	-	11%	20%	69%
P4vir1	11.5Kb	-	-	20%
P4vir1del10	10.5Kb	-	-	95%

The origin of these knotted structures is not well understood, nevertheless the knot spectrum is a marker for DNA packing in the viral capsid.

EFFECTS OF CONFINEMENT ON DNA KNOTTING

- No confinement--3% knots, mostly trefoils
- Viral knots--95% knots, very high complexity--average crossover number 27!

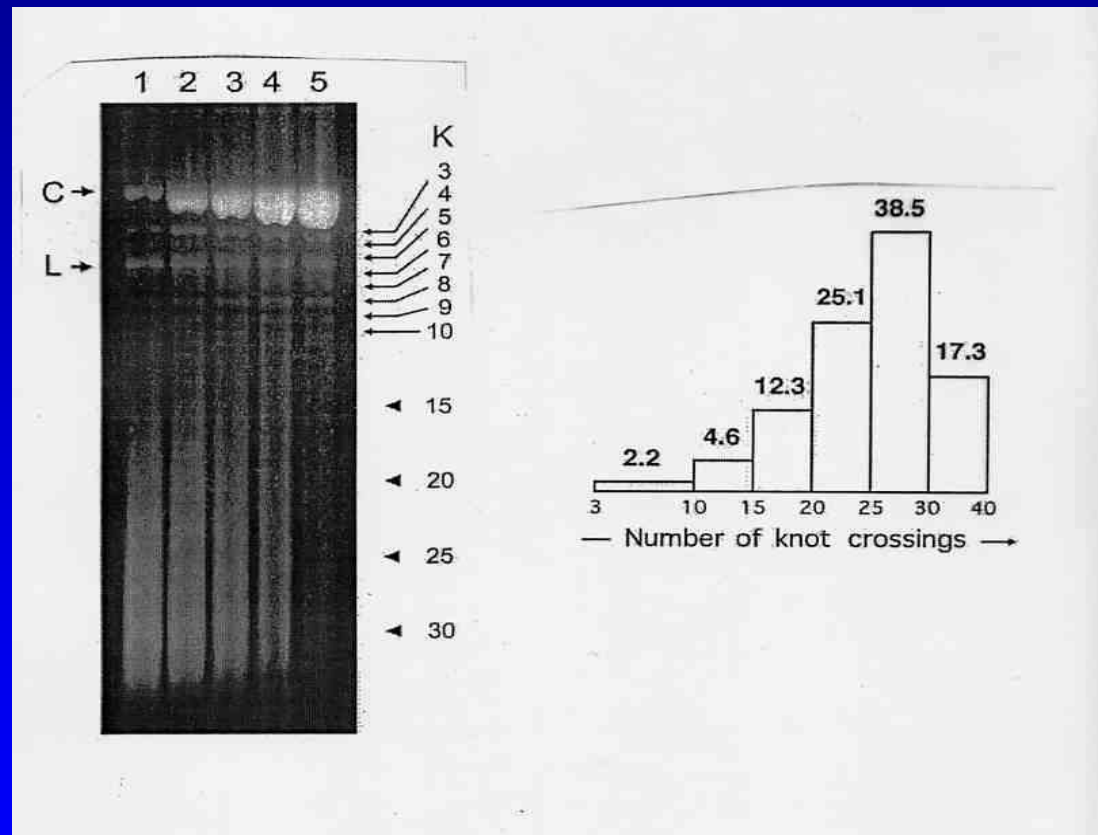
MATURE vs TAILLESS PHAGE



Mutants--48% of knots formed inside capsid

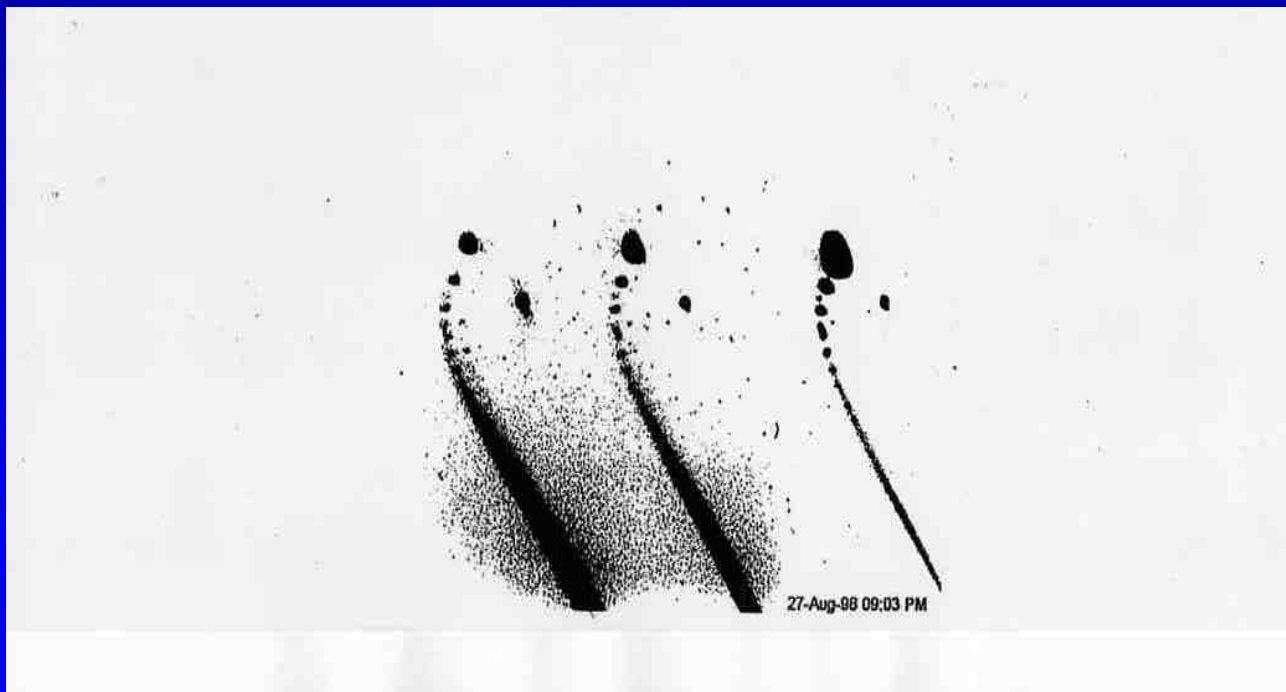
Arsuaga et al, PNAS 99 (2002), 5373

P4 KNOT SPECTRUM

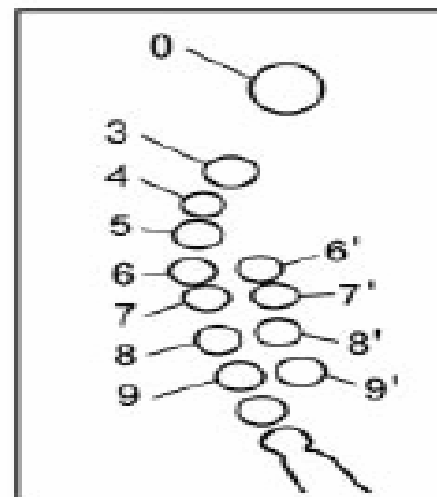
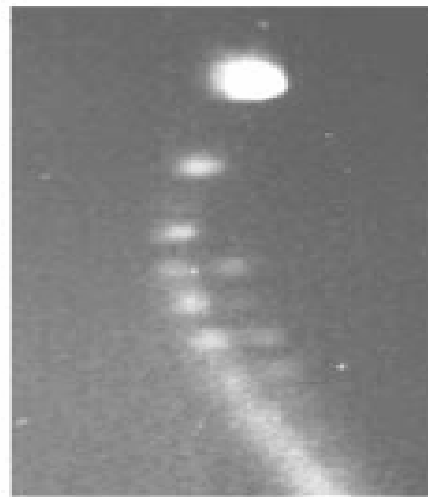


97% of DNA knots had crossing number > 10!
Arsuaga et al, PNAS 99 (2002), 5373

2D GEL



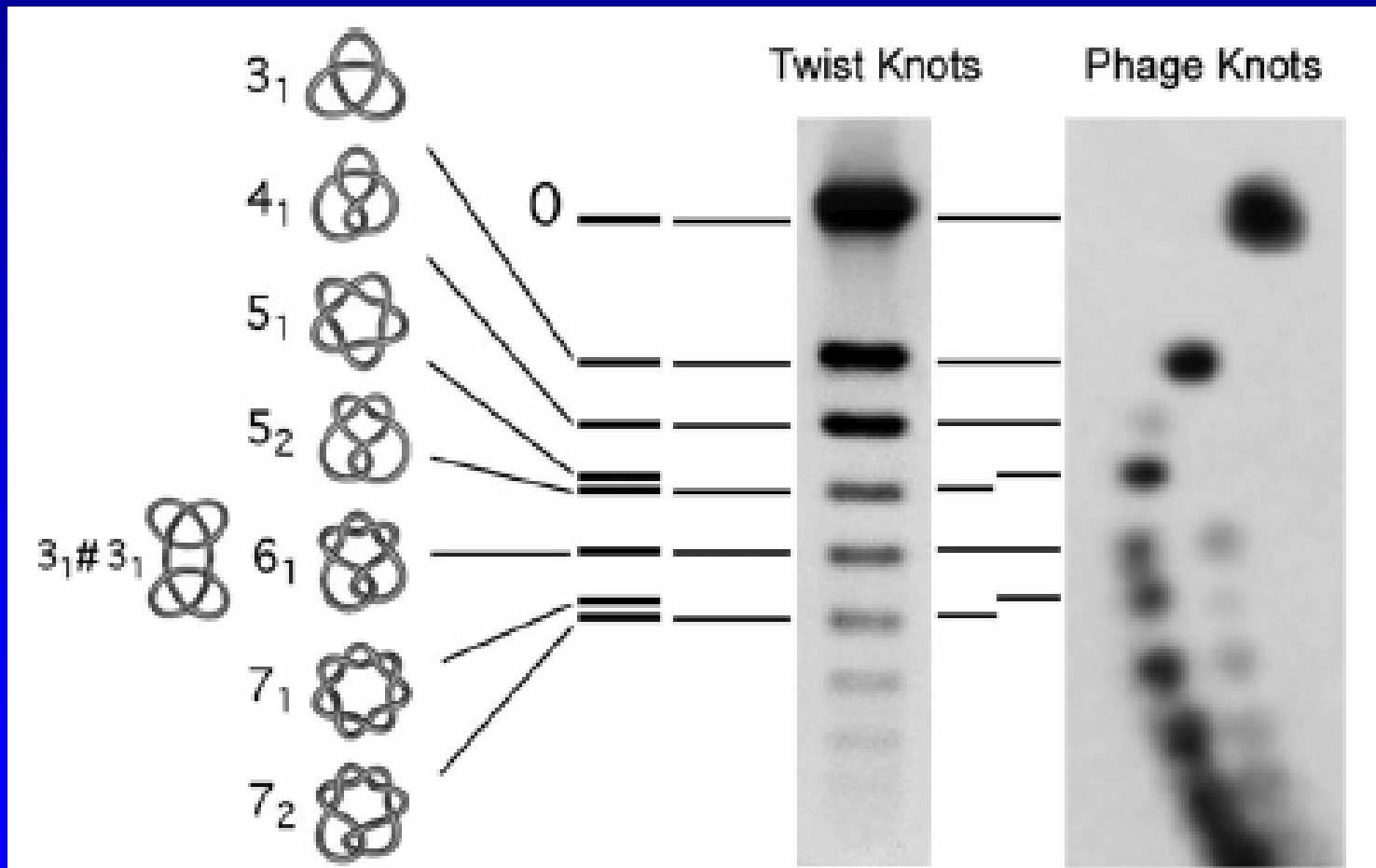
2D GEL RESOLVES SMALL KNOTS



Total DNA		Individual Knot populations	
Linear	3 %	3 : 0.36 %	
Unknotted	2 %	4 : 0.02 %	
Knotted	95 %	5 : 0.26 %	
		6 : 0.13 %	6' : 0.08 %
		7 : 0.27 %	7' : 0.03 %
		8 : 0.34 %	8' : 0.10 %
		9 : 0.39 %	9' : 0.09 %

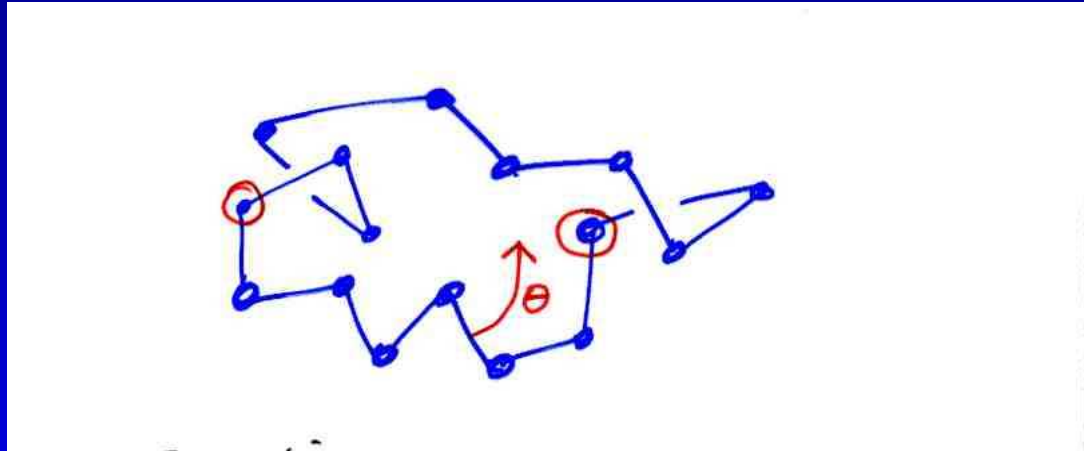
Arsuaga et al, PNAS 102 (2005), 9165

2D GEL RESOLVES SMALL KNOTS



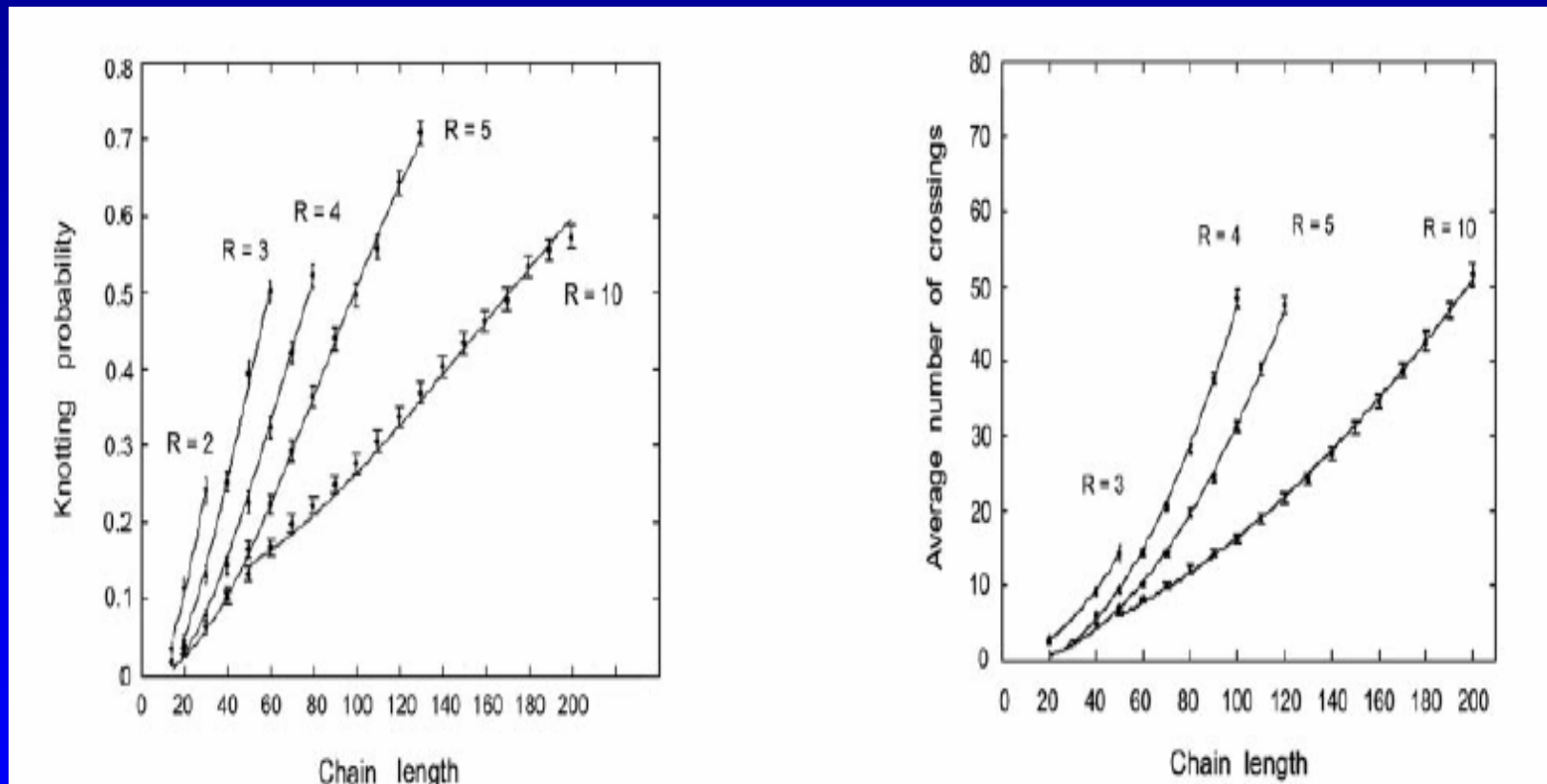
Arsuaga et al, PNAS 102 (2005), 9165

PIVOT ALGORITHM



- Ergodic--no volume exclusion in our simulation
- $\Delta(-1)$, $\Delta(-2)$, $\Delta(-3)$ as knot detector
- Space filling polymers in confined volumes--very difficult to simulate

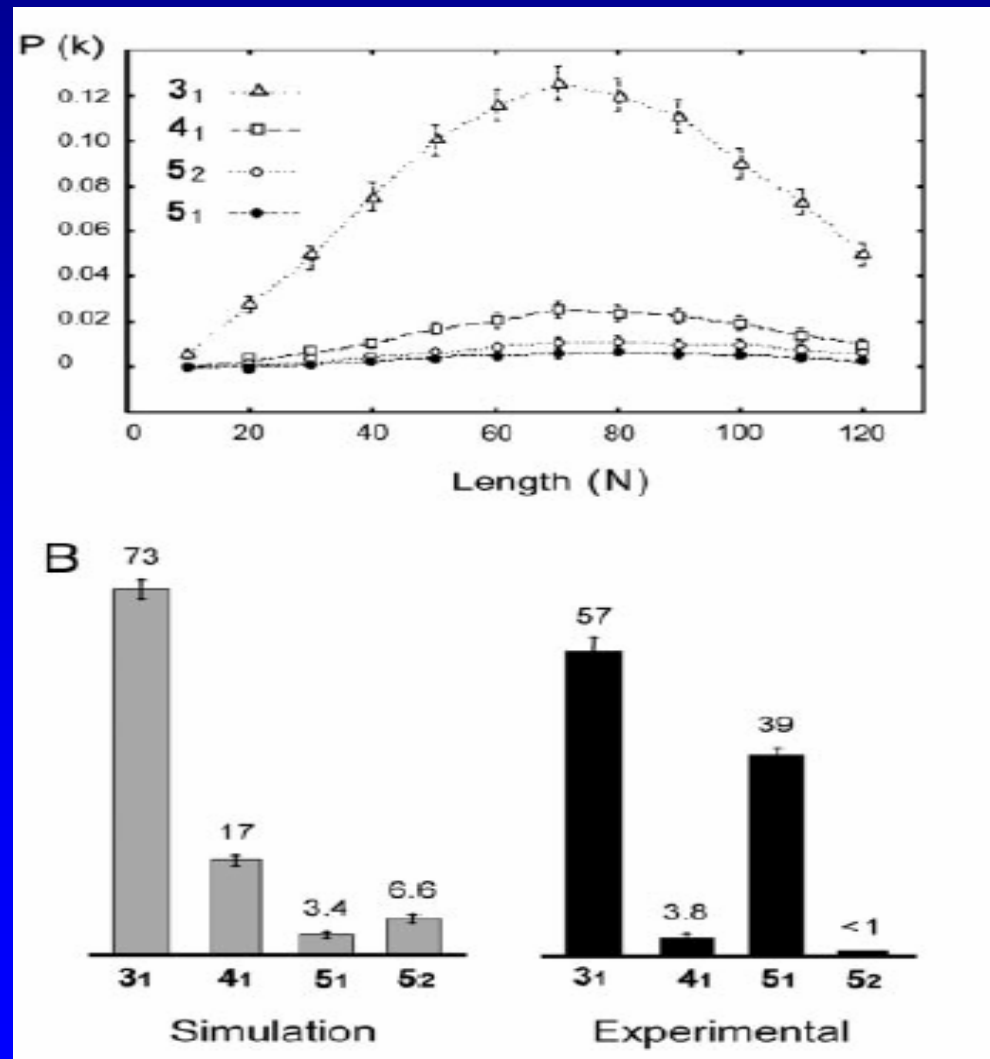
VOLUME EFFECTS ON KNOT SIMULATION



- On average, 75% of crossings are extraneous

Arsuaga et al, PNAS 99 (2002), 5373

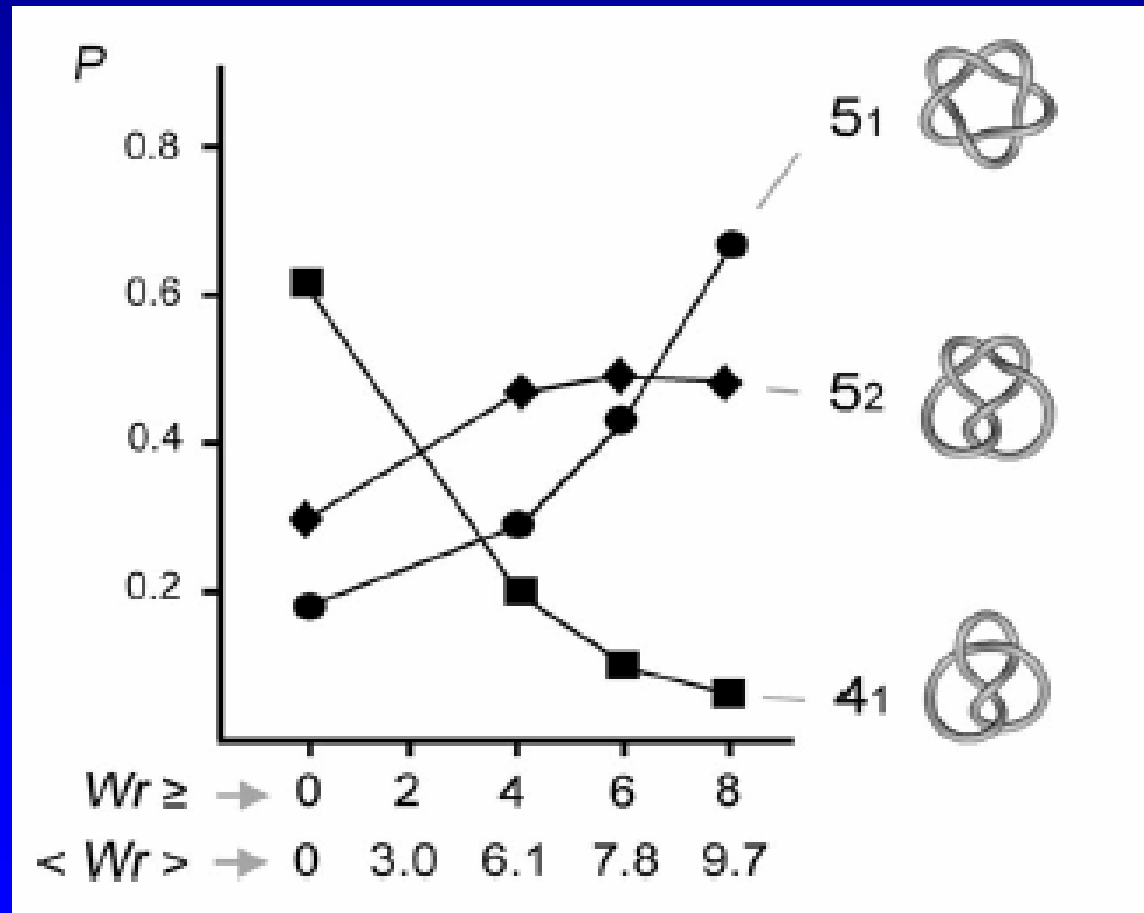
SIMULATION vs EXPERIMENT



$n=90, R=4$

Arsuaga et al, PNAS 102 (2005), 9165

EFFECT OF WRITHE-BIASED SAMPLING

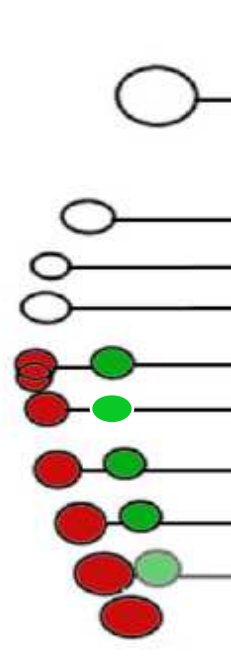


$n=90, R=4$

Arsuaga et al, PNAS 102 (2005), 9165

Consistent with a high **writhe** of the packed DNA:

- Low amount of knot 4_1 and of composite $3_1 \# 4_1$
- Predominance of torus knots (elevated writhe)



3_1 $Wr = 3.41$

4_1 $Wr = 0.00$

5_1 $Wr = 6.26$

5_2 $Wr = 4.54$

6_1 $Wr = 1.23$

6_2 $Wr = 2.70$

6_3 $Wr = 0.16$

7_1 $Wr = 9.15$

$3_1 \# 3_1$ $Wr = 6.81$

$3_1 \# -3_1$ $Wr = 0.01$

$3_1 \# 4_1$ $Wr \sim 3.41$

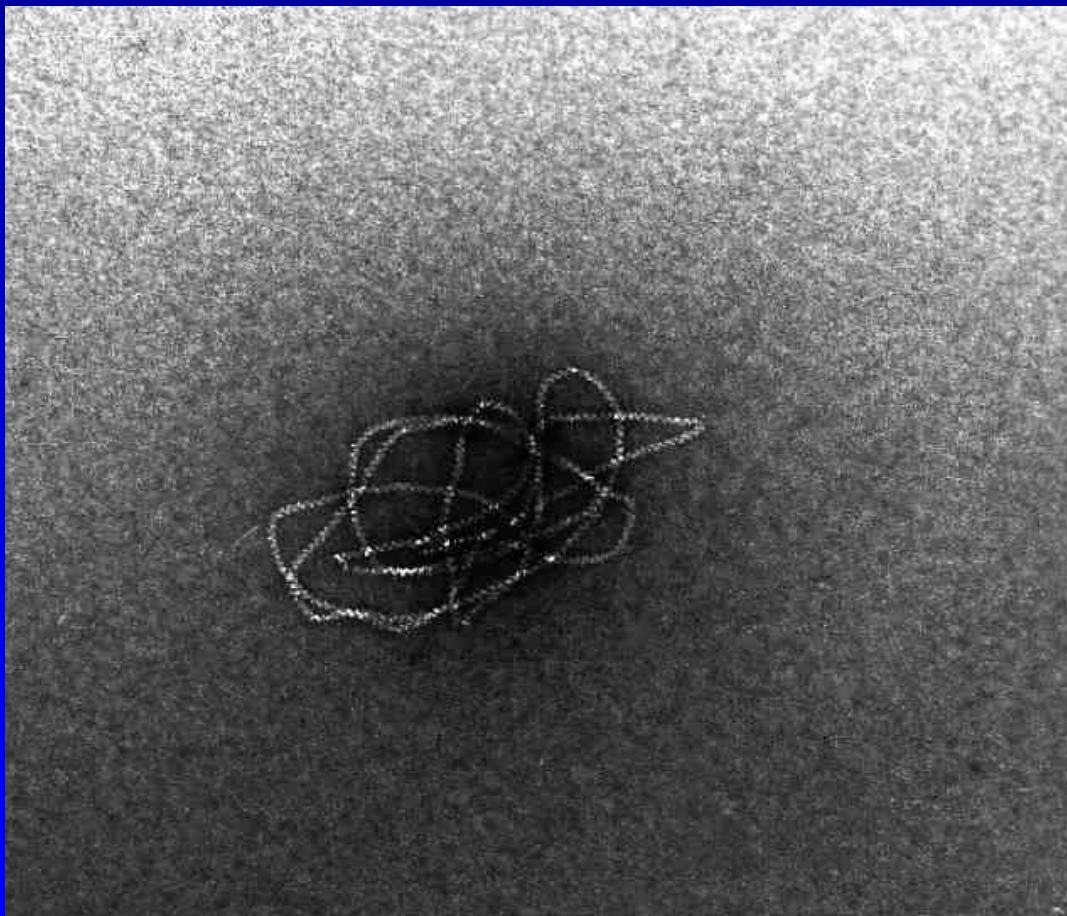
primes

composites

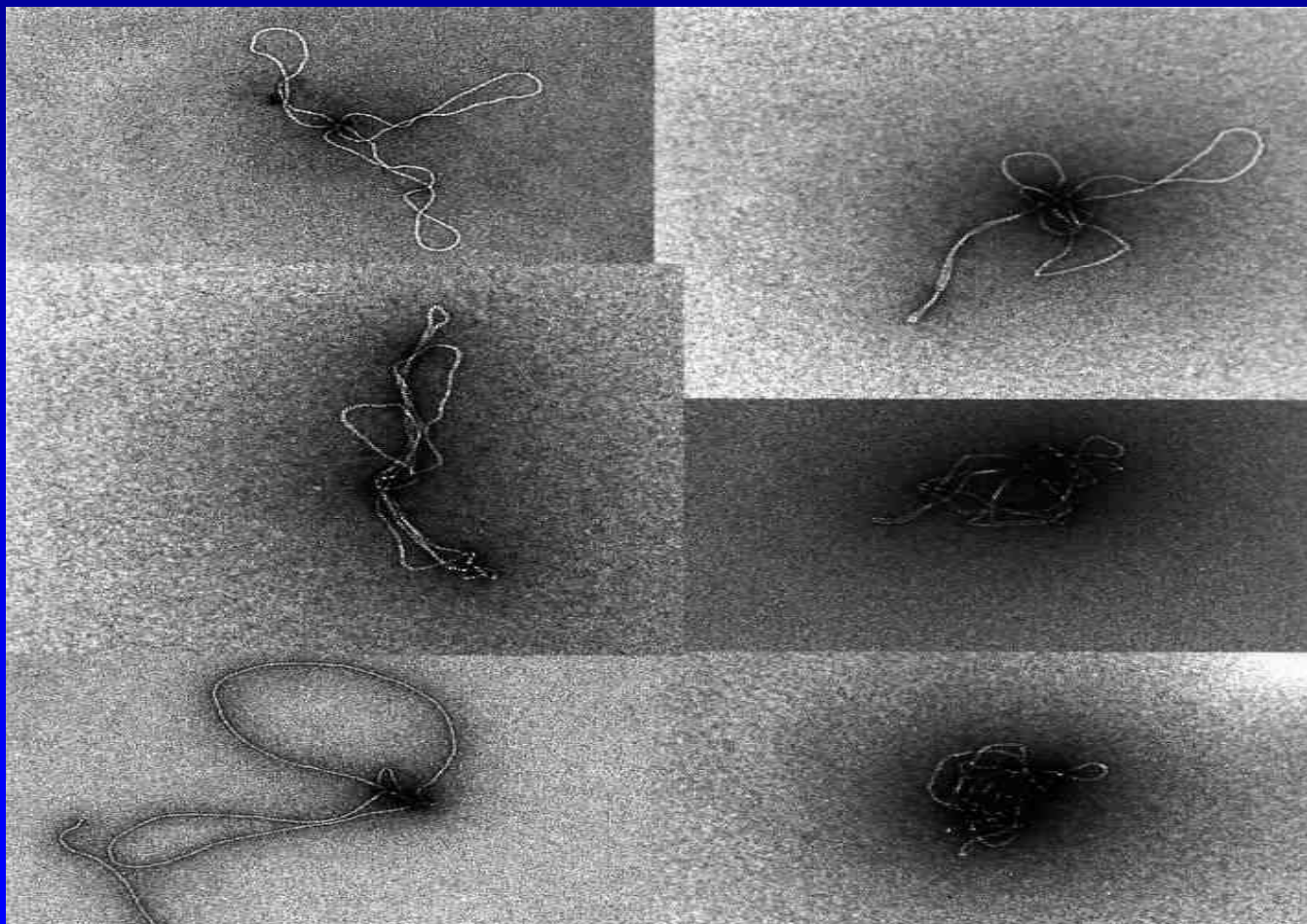
CONCLUSIONS

- Viral DNA not randomly embedded (4_1 and 5_2 deficit, 5_1 and 7_1 excess in observed knot spectrum)
- Viral DNA has a chiral packing mechanism--writhe-biased simulation close to observed spectrum
- Torus knot excess favors toroidal or spool-like packing conformation of capsid DNA
- Next step--EM (AFM) of 3- and 5- crossing knots to see if they all have same chirality

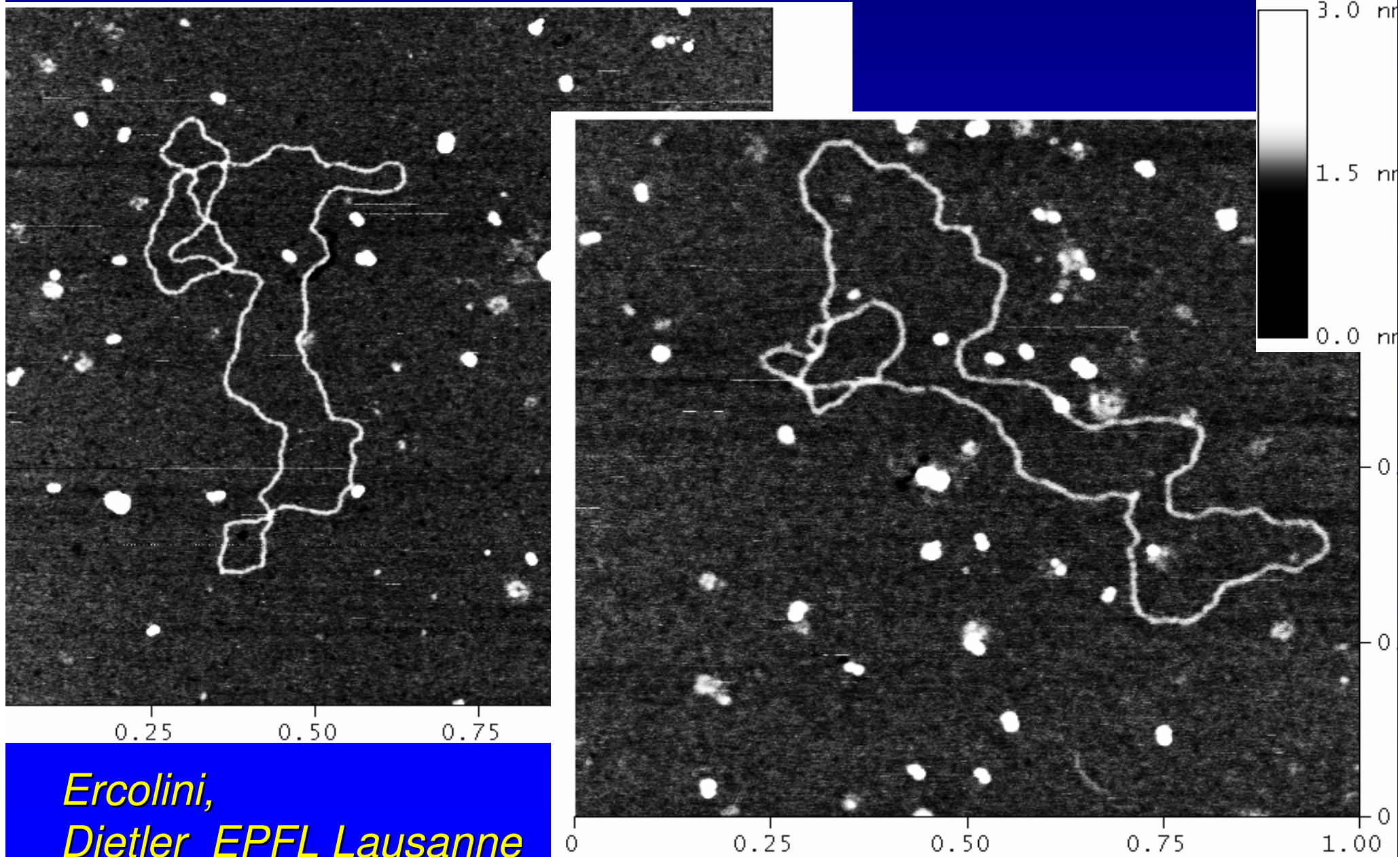
UNKNOWN P4 KNOT



UNKNOWN P4 KNOTS



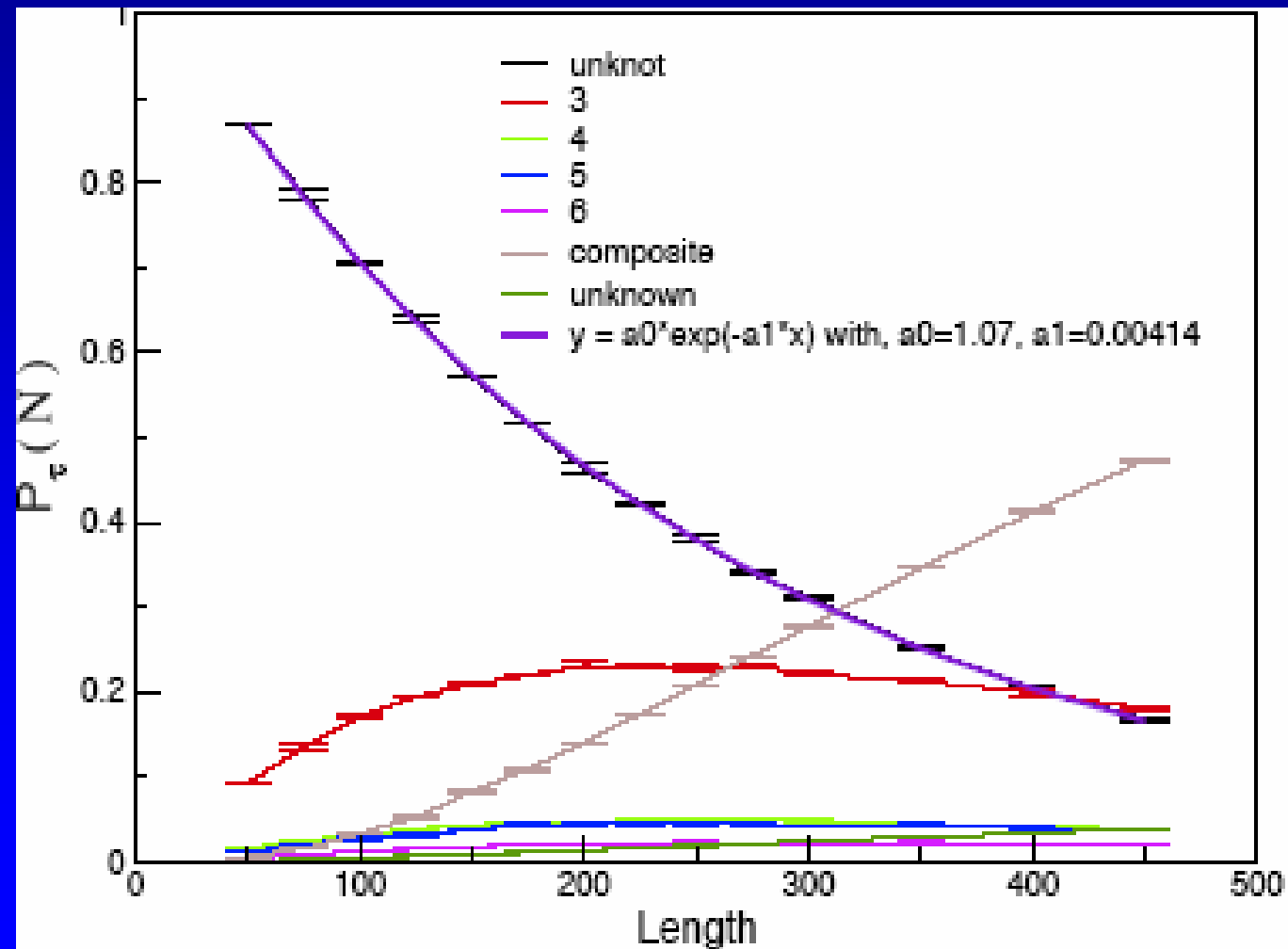
AFM Images of Simple DNA Knots (Mg^{2+})



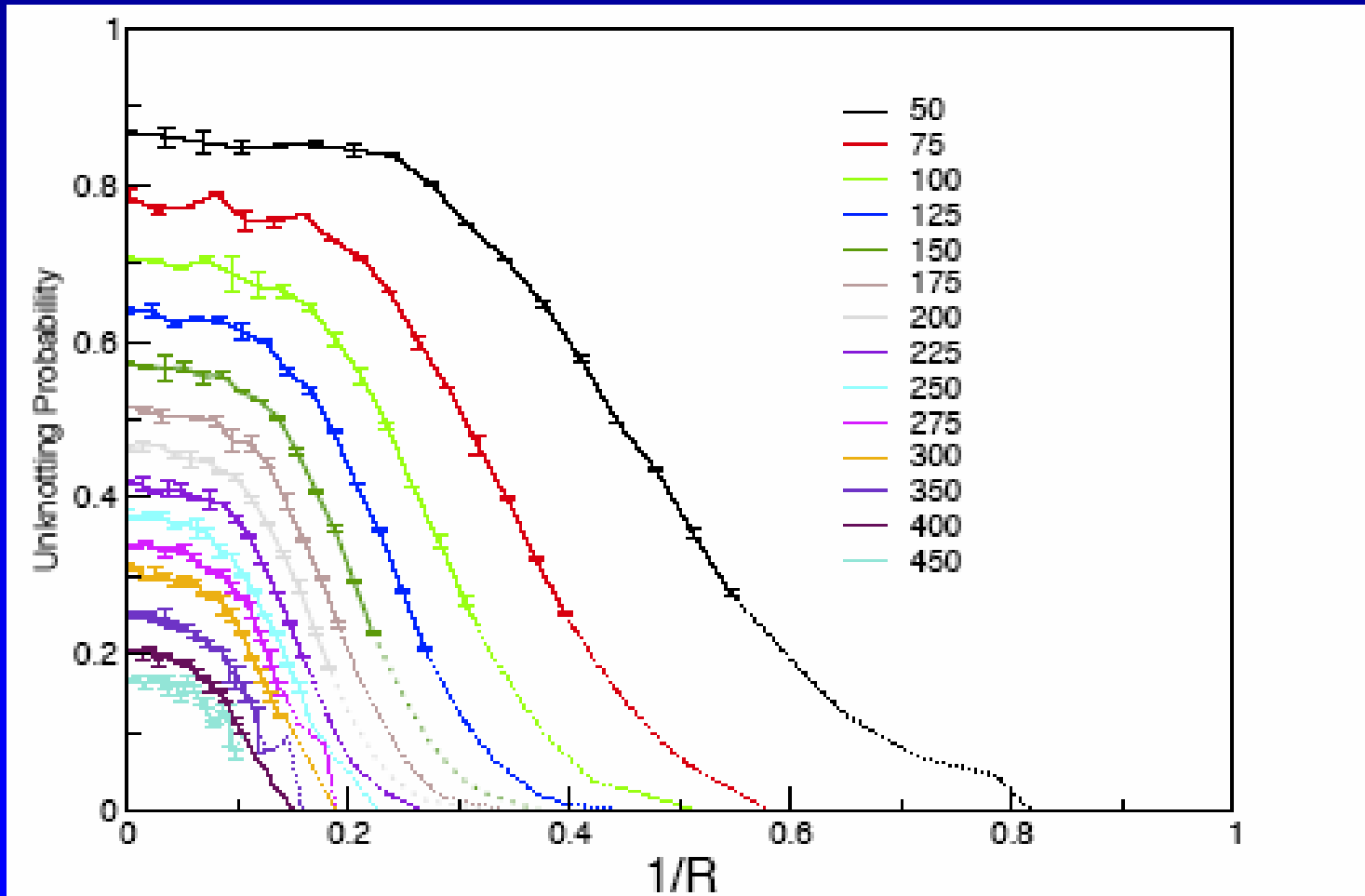
NEW SIMULATION

- Parallel tempering scheme
- Smooth configuration to remove extraneous crossings
- Use KnotFind to identify the knot-ID's prime and composite knots of up to 16 crossings
- Problem--some knots cannot be ID'd--might be complicated unknots!

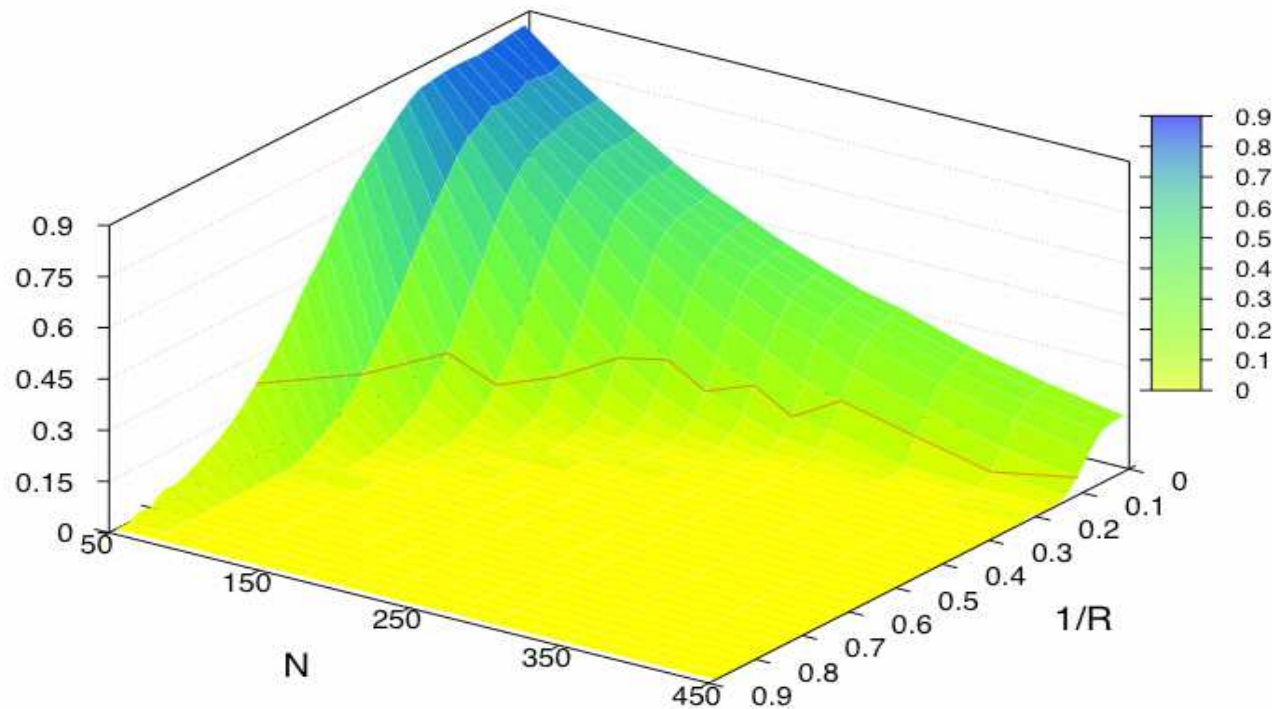
UNCONSTRAINED KNOTTING PROBABILITIES



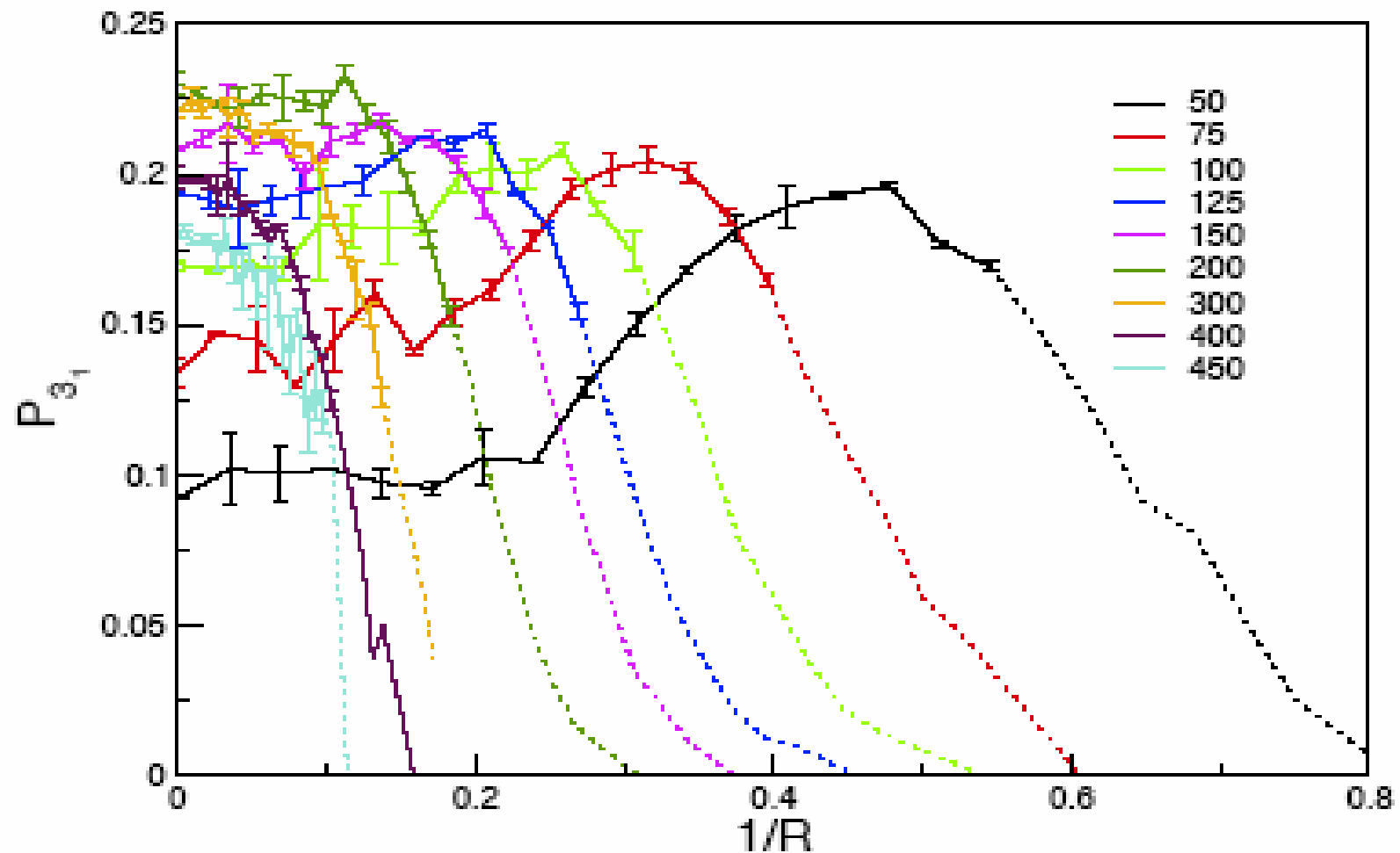
CONSTRAINED UNKNOTTING PROBABILITY



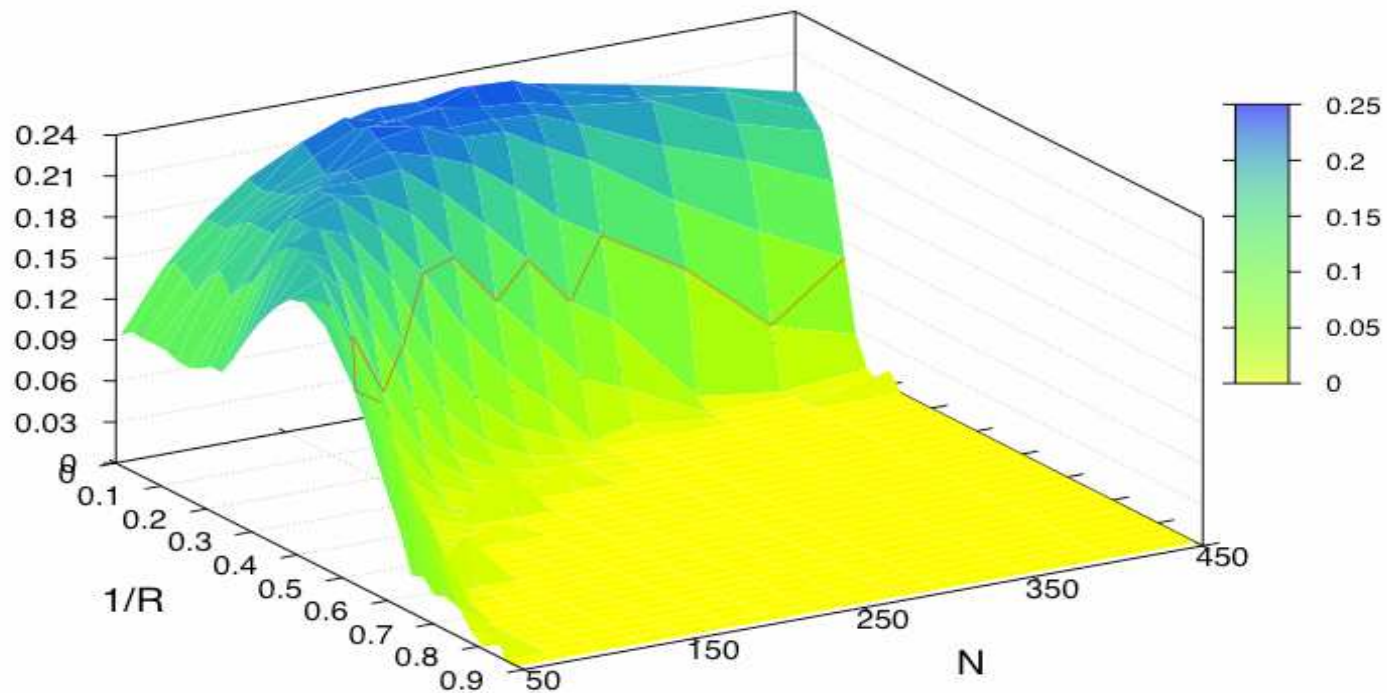
CONSTRAINED UNKNOTTING PROBABILITY



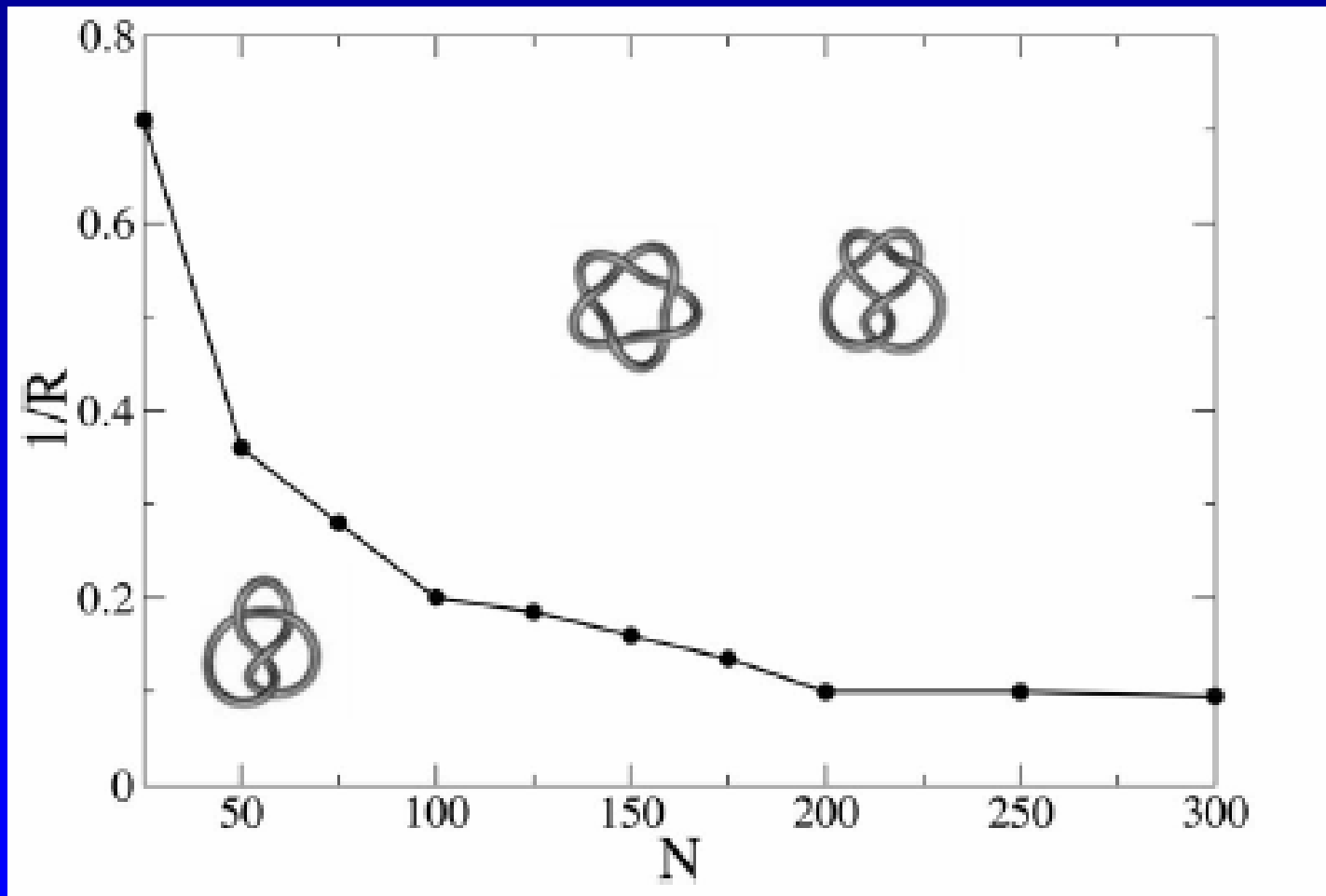
CONSTRAINED TREFOIL KNOT PROBABILITIES



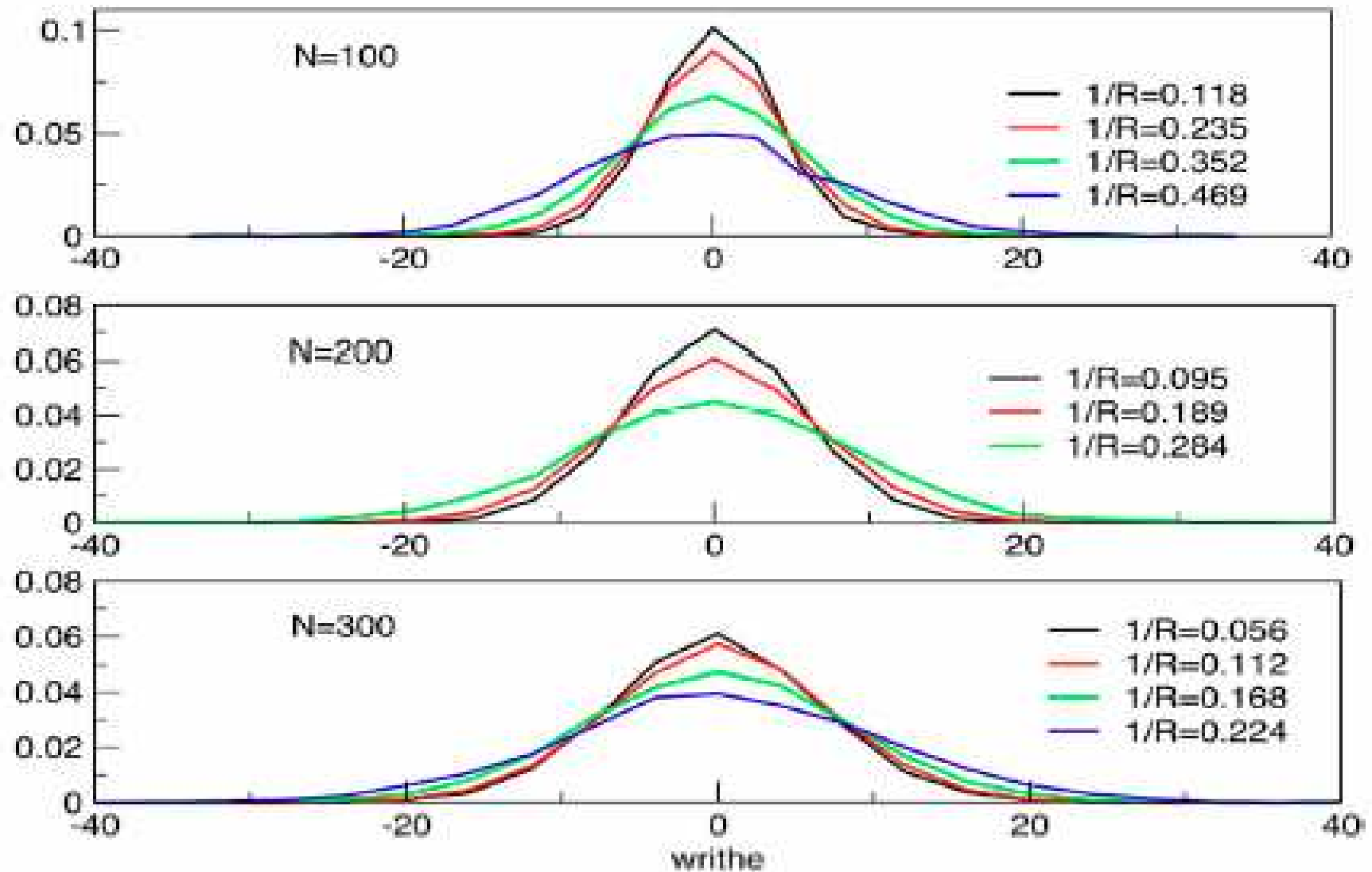
CONSTRAINED TREFOIL PROBABILITY



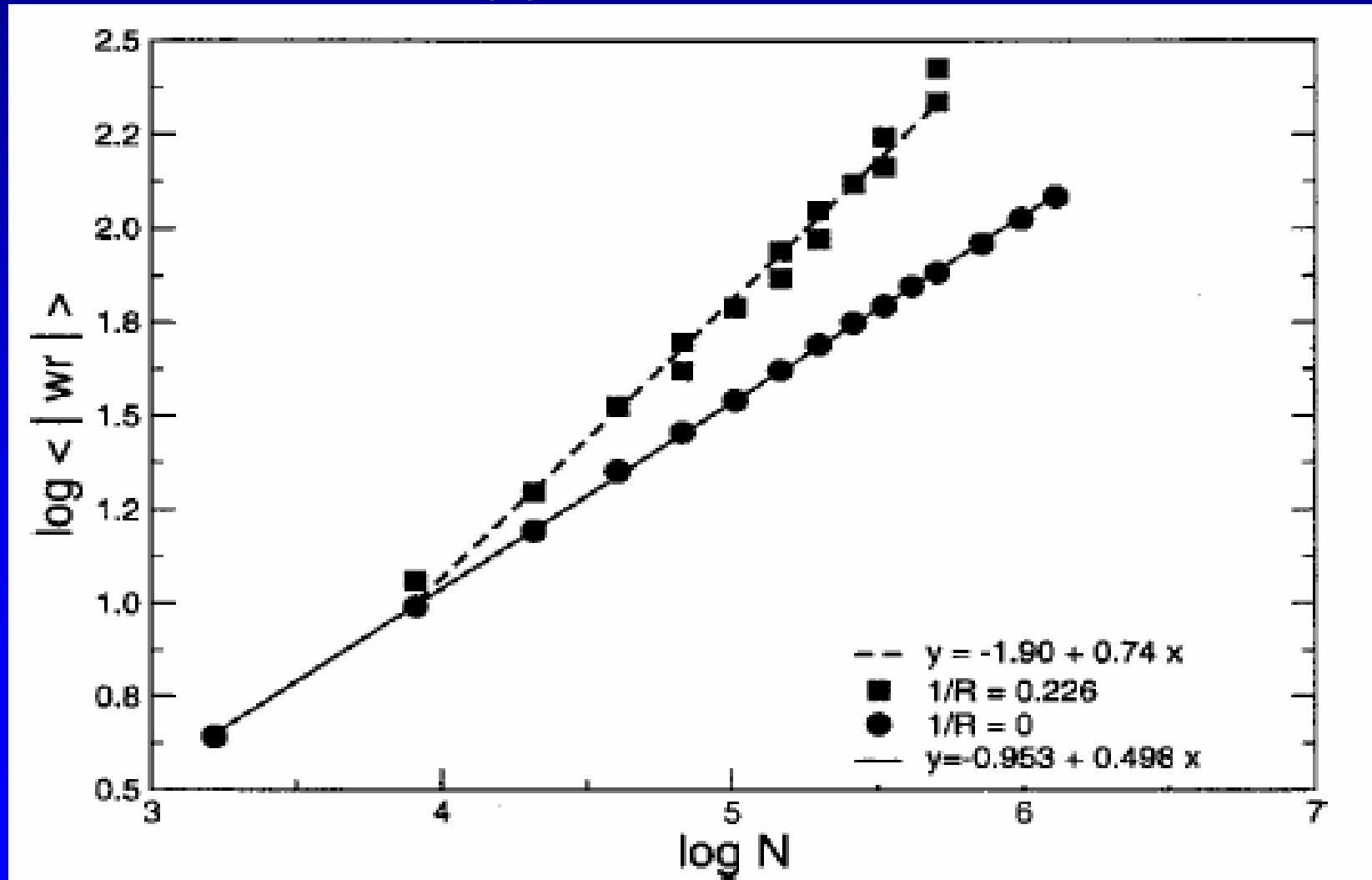
4 vs 5 CROSSING PHASE DIAGRAM



CONFINED WRITHE



GROWTH OF CONFINED WRITHE



COLLABORATORS

- Stu Whittington
- Buks van Rensburg
- Carla Tesi
- Enzo Orlandini
- Chris Soteros
- Yuanan Diao
- Nick Pippenger
- Javier Arsuaga
- Mariel Vazquez
- Joaquim Roca
- P. McGuirk
- Christian Micheletti
- Davide Marenduzzo

Thank You

- National Science Foundation
- Burroughs Wellcome Fund