

Untangling **Knots** in **Lattices** and **Proteins**

A Computational Study

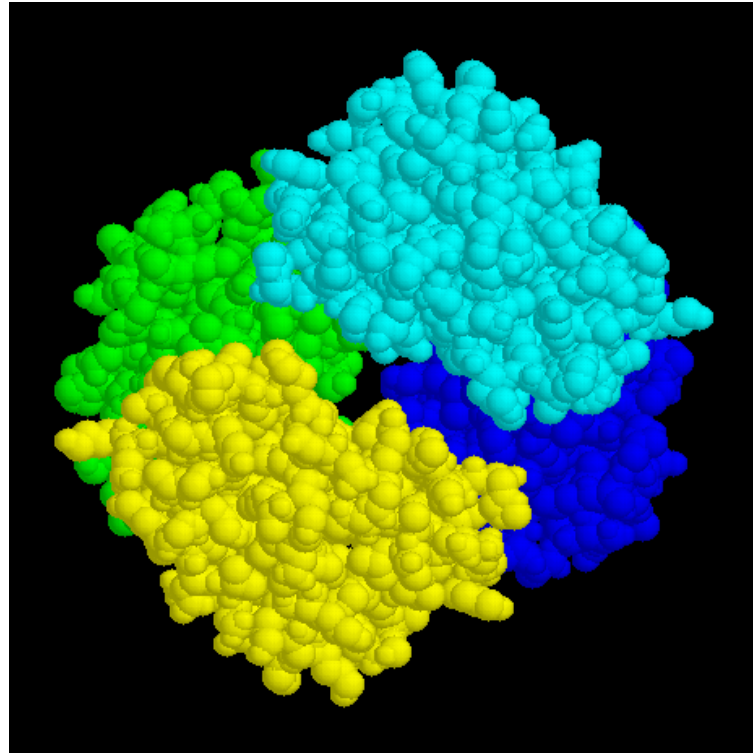
By **Rhonald Lua**

Adviser: **Alexander Yu. Grosberg**

University of Minnesota



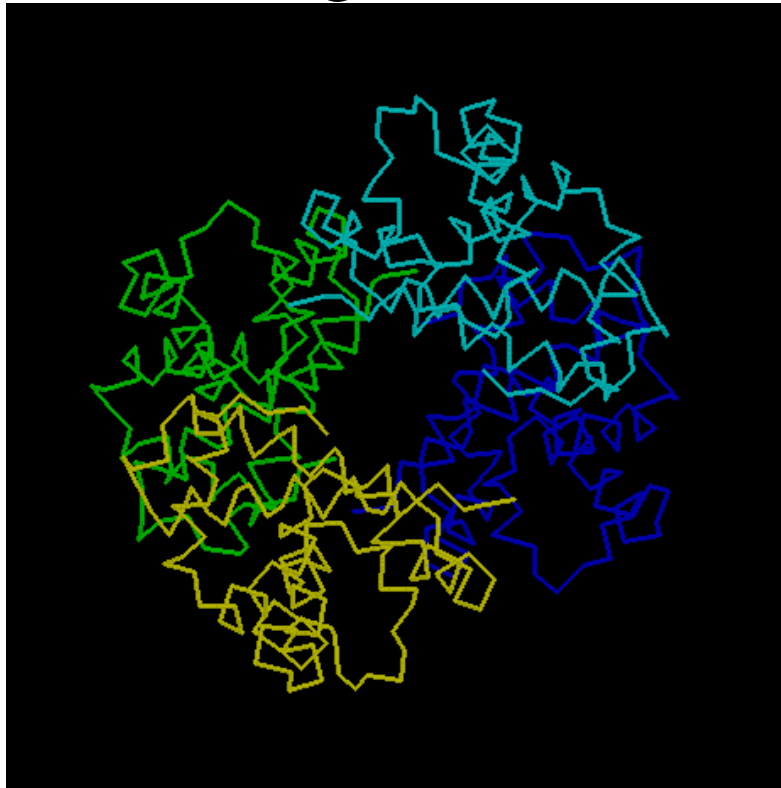
Human Hemoglobin (oxygen transport protein)



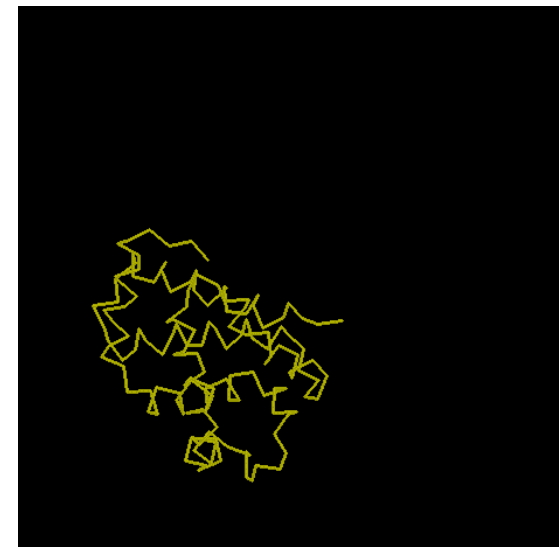
(Structure by
G. FERMI and
M.F. PERUTZ)

Globular proteins have dense, crystal-like packing density.
Proteins are small biomolecular “machines” responsible
for carrying out many life processes.

Hemoglobin Protein Backbone (string of α -carbon units)

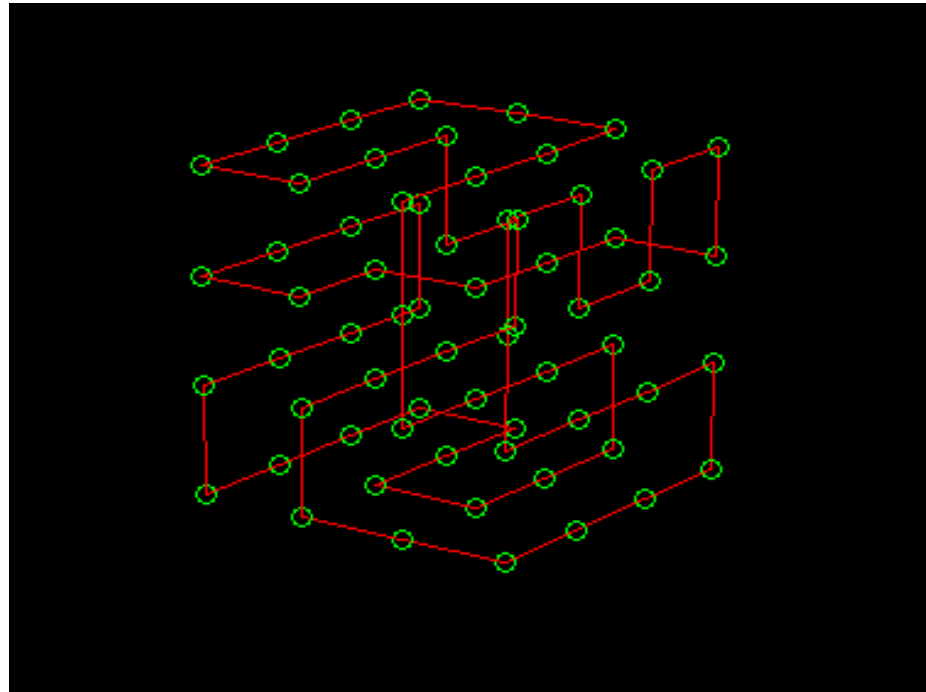


One chain



“ball of yarn”

4x4x4 Compact Lattice Loop



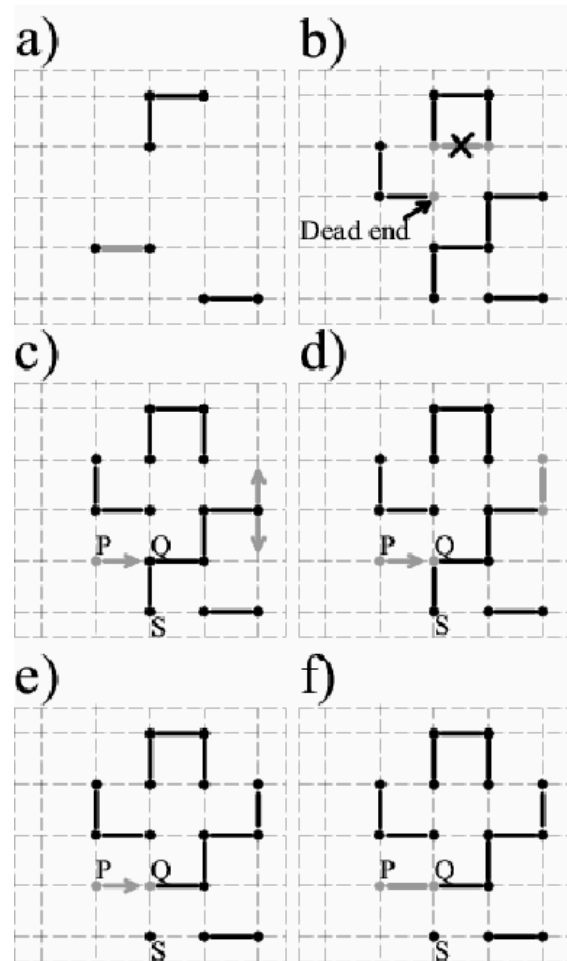
Possible cube dimensions: $2 \times 2 \times 2, 4 \times 4 \times 4, 6 \times 6 \times 6, \dots, L \times L \times L, \dots$

No. of distinct conformations: $\left(\frac{z-1}{e}\right)^N$ (Flory)

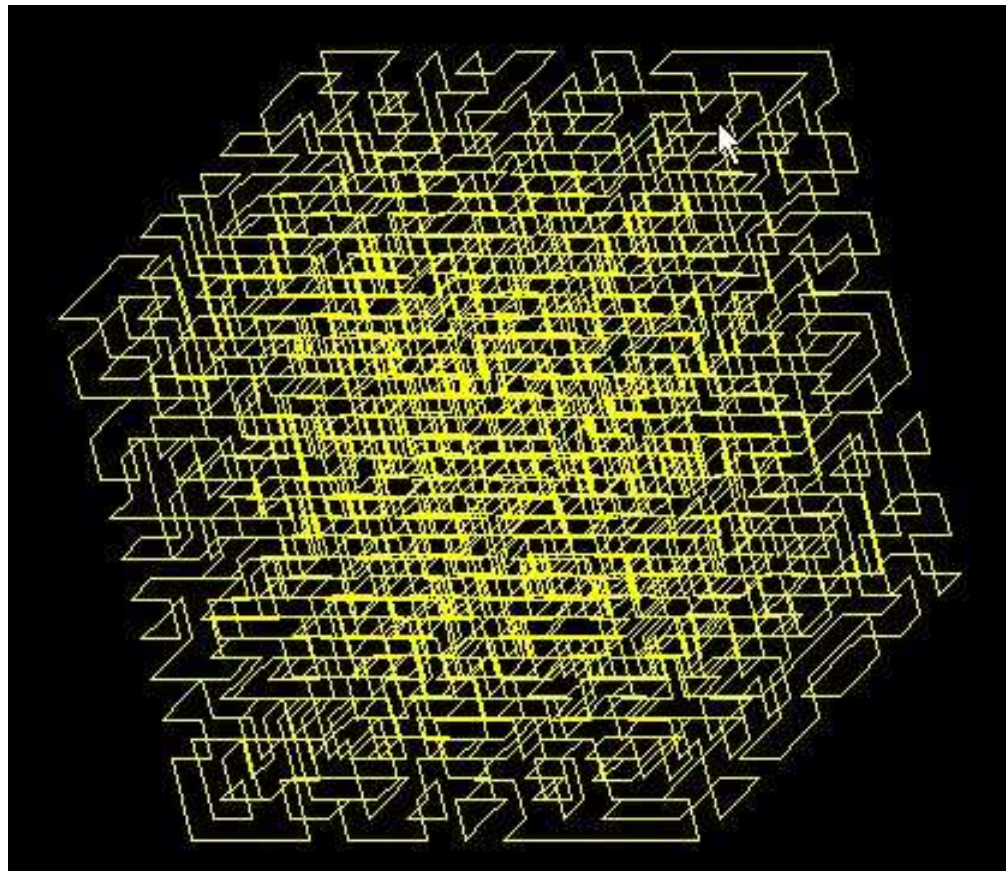
$z = 6$ in 3D

Hamiltonian Path Generation

(A. Borovinskiy, based on work by R. Ramakrishnan, J.F. Pekny, J.M. Caruthers)



14x14x14 Compact Lattice Loop

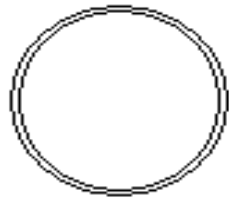


In this talk...

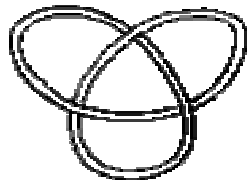
- knots and their relevance to physics
- “virtual” tools to study knots
- knotting probability of compact lattice loops
- statistics of subchains in compact lattice loops
- knots in proteins

Knot – a closed curve in space that does not intersect itself.

The first few knots:



Trivial knot (Unknot)
0-1



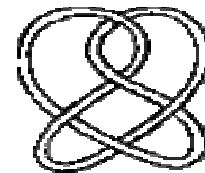
3-1 (Trefoil)



4-1 (Figure-8)



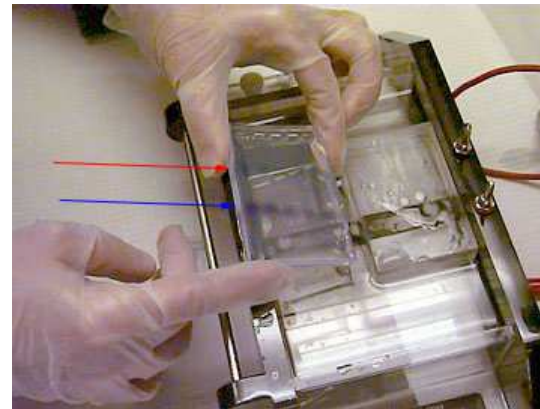
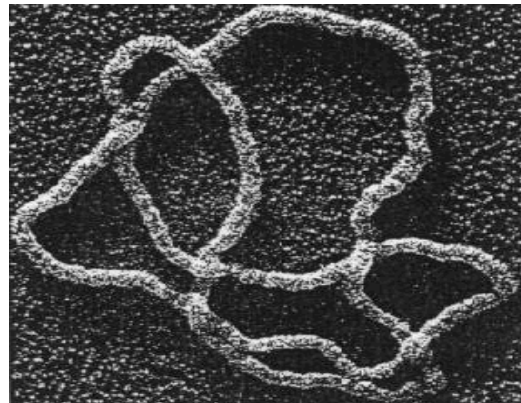
5-1 (Cinquefoil, Pentafoil
Solomon's seal)



5-2

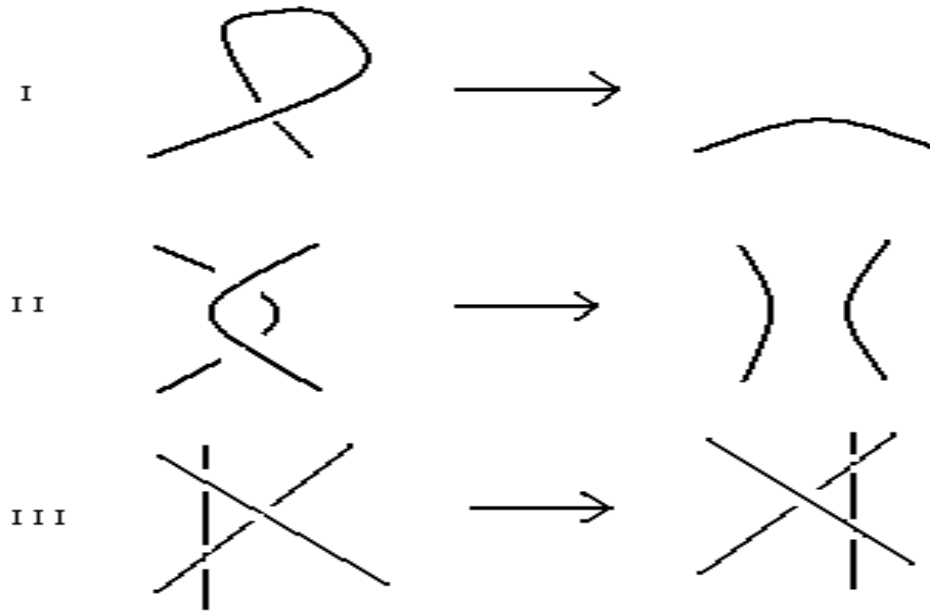
Knots in Physics

- Lord Kelvin (1867): Atoms are knots (vortices) of some medium (ether).
- Knots appear in Quantum Field Theory and Statistical Physics.
- Knots in biomolecules. Example: The more complicated the knot in circular DNA the faster it moves in gel-electrophoresis experiments



A Little Knot Math

Reidemeister Moves

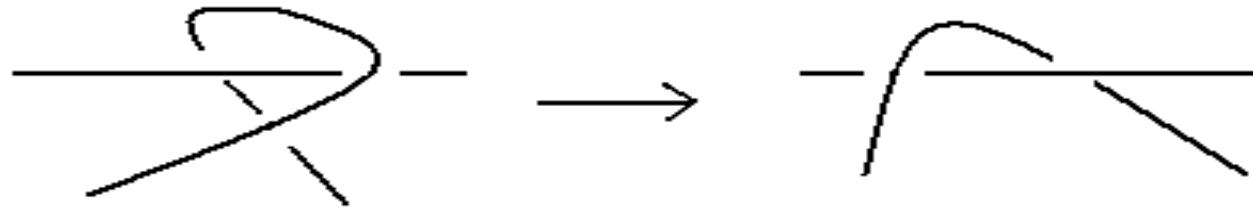


Reidemeister's Theorem:

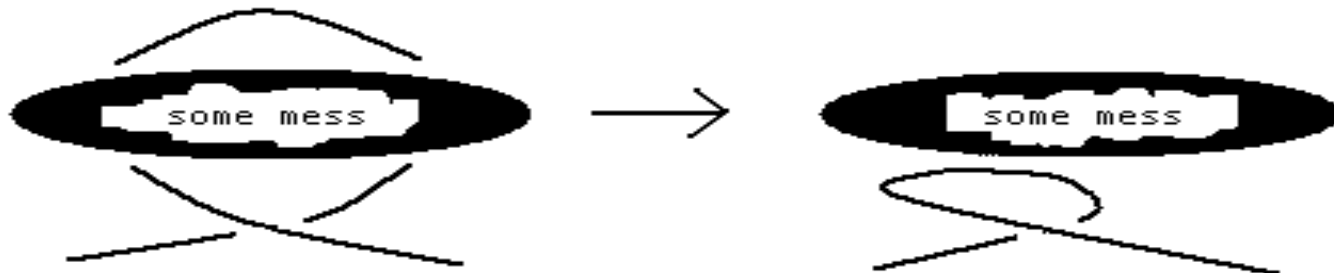
Two knots are equivalent if and only if any diagram of one may be transformed into the other via a sequence of Reidemeister moves.

Compounded Reidemeister Moves

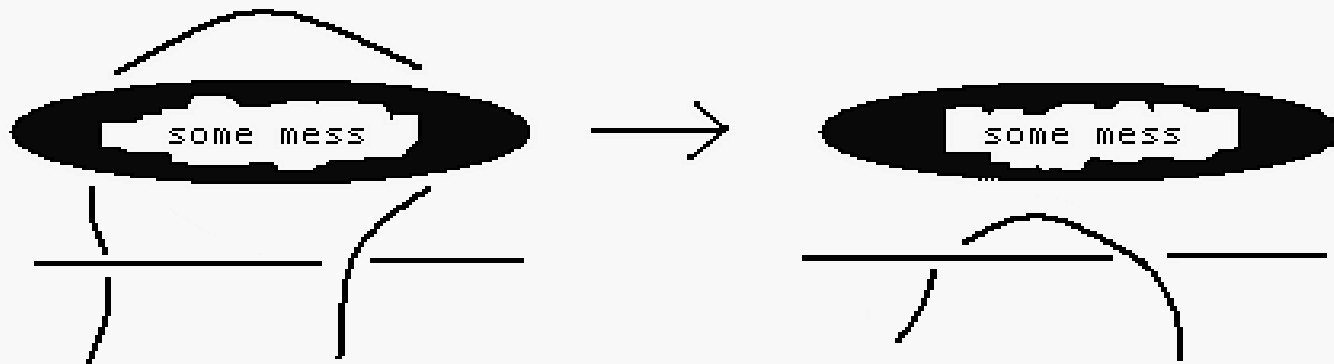
I+III



macro
move



macro
move

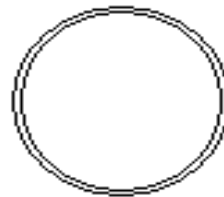


Knot Invariants -Mathematical signatures of a knot.

Name	Symbol
Alexander	$\Delta(-1)$
Vassiliev degree 2	v_2
Vassiliev degree 3	v_3

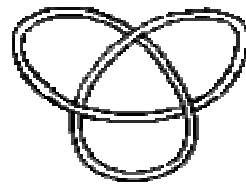
Examples:

Trivial knot
0-1



$\Delta(-1)=1$
 $v_2=0$
 $v_3=0$

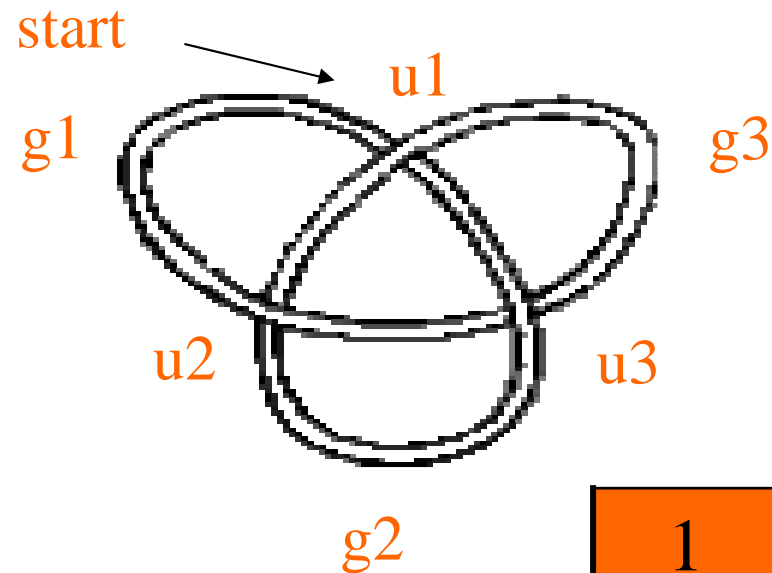
Trefoil knot
3-1



$\Delta(-1)=3$
 $v_2=1$
 $v_3=1$

Alexander Polynomial, $\Delta(t)$

(first knot invariant/signature)



Alexander matrix for this trefoil:

1	-t	t-1
t-1	1	-t
0	t-1	1

Alexander invariant:

$$\Delta(-1) = \det$$

1	1
-2	1

$$= 3$$

Recipe for Constructing Alexander Matrix, a_{kj}

$n \times n$ matrix where n is
the number of underpasses

In the following
index k corresponds to k th underpass and
index i corresponds to the generator number
of the arc overpassing the k th underpass

For row k :

1) when $i=k$ or $i=k+1$ then

$$a_{kk} = -1, a_{kk+1} = 1$$

2) when i equals neither k nor $k+1$:

If the crossing has sign -1 :

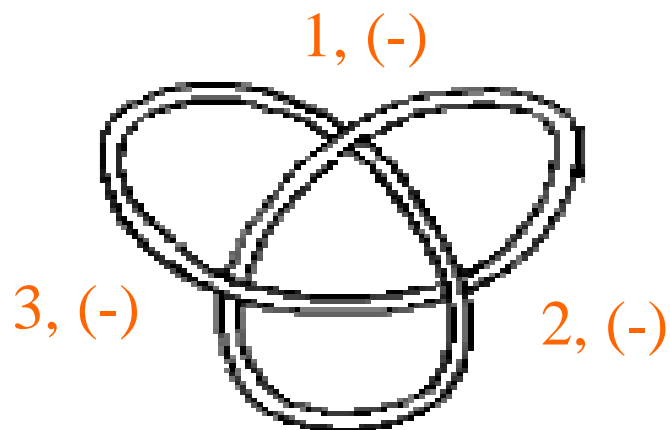
$$a_{kk} = 1, a_{kk+1} = -t, a_{ki} = t-1$$

If the crossing has sign $+1$:

$$a_{kk} = -t, a_{kk+1} = 1, a_{ki} = t-1$$

3) All other elements are zero.

Gauss Code and Gauss Diagram



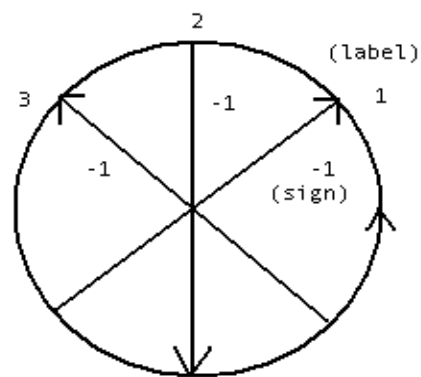
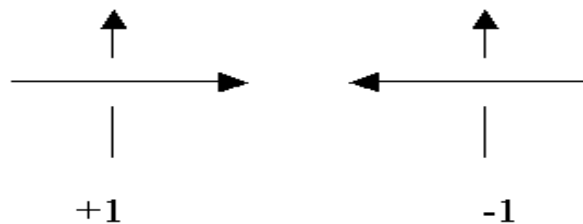
Gauss code for left-handed trefoil:
 $b - 1, a - 2, b - 3, a - 1, b - 2, a - 3$

(Alternatively...)

Gauss Diagram for trefoil:

a – ‘above’
 b – ‘below’

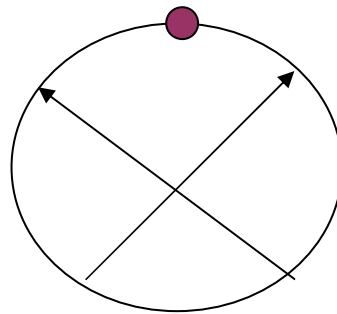
sign:



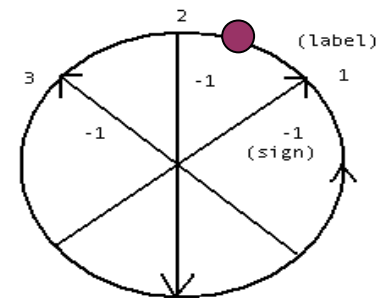
Vassiliev Invariants

(Diagram methods by M. Polyak and O. Viro)

Degree two (v_2): Look for this pattern:



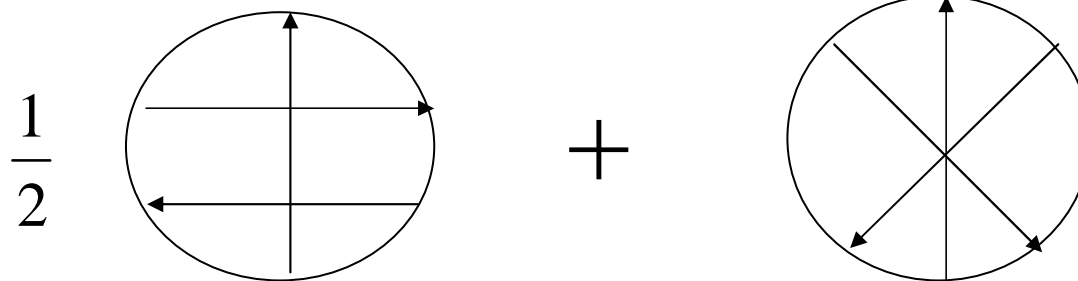
e.g. trefoil



$$v_2 = 1$$

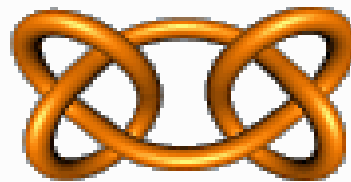
$$v_3 = 1$$

Degree three (v_3): Look for these patterns:

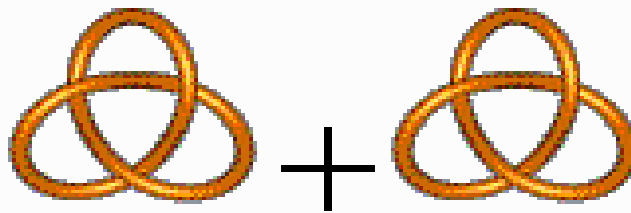


Prime and Composite Knots

Composite knot, K



K_1



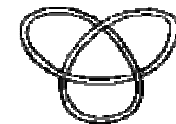
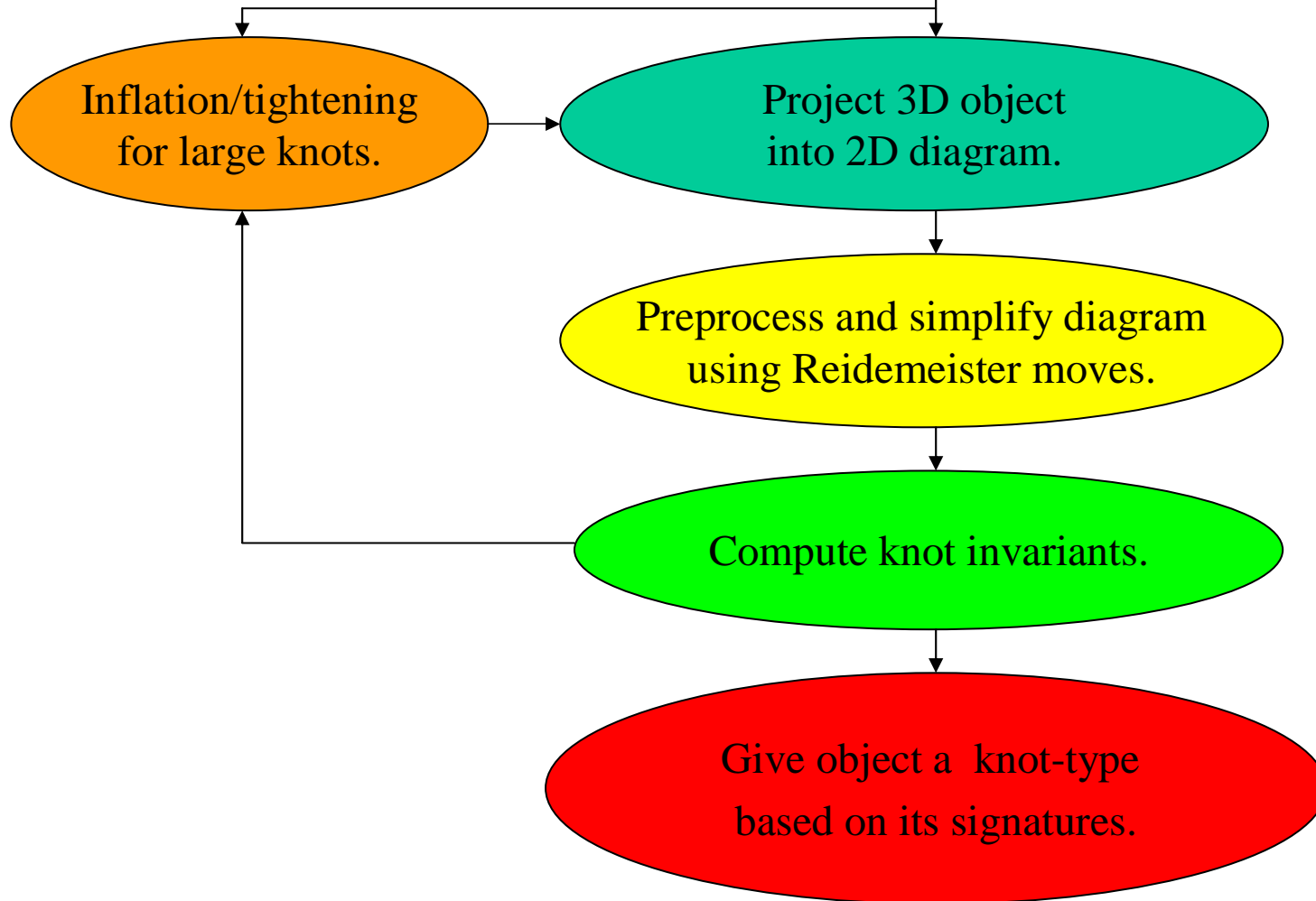
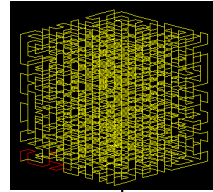
K_2

Alexander: $\Delta(t)_K = \Delta(t)_{K_1} \Delta(t)_{K_2}$

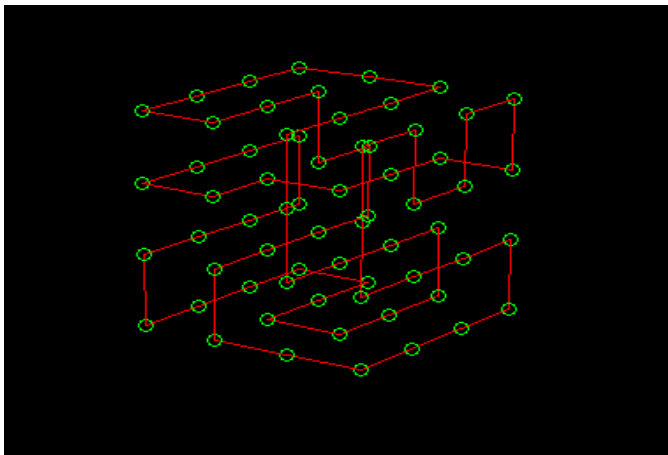
Vassiliev: $v_2(K) = v_2(K_1) + v_2(K_2)$

$v_3(K) = v_3(K_1) + v_3(K_2)$

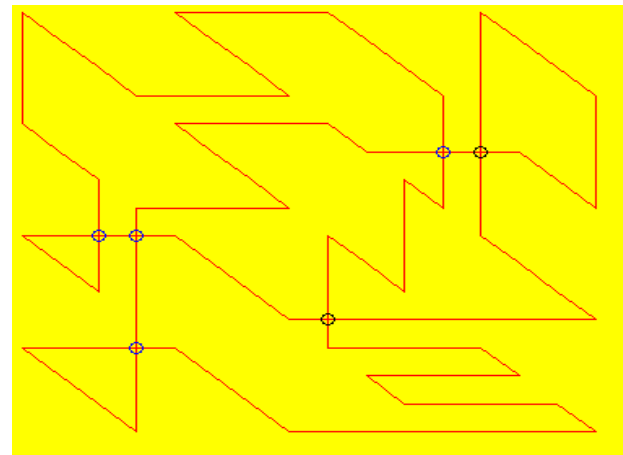
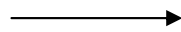
Method to Determine Type of Knot



A. Projection

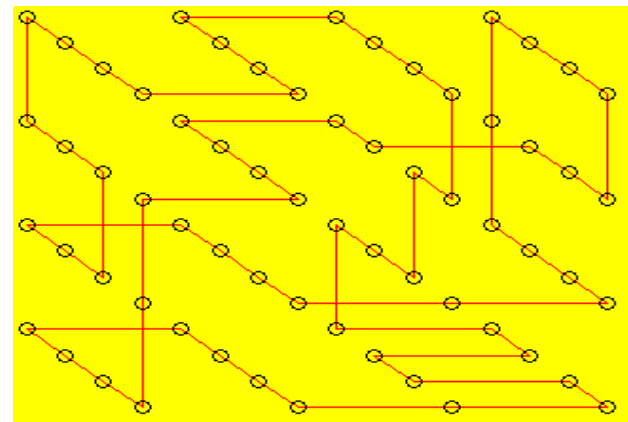
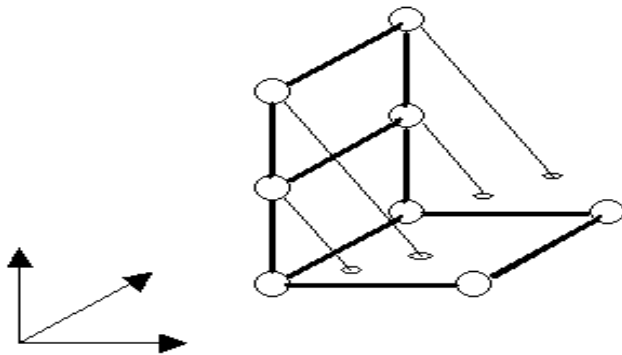


3D conformation



2D knot projection

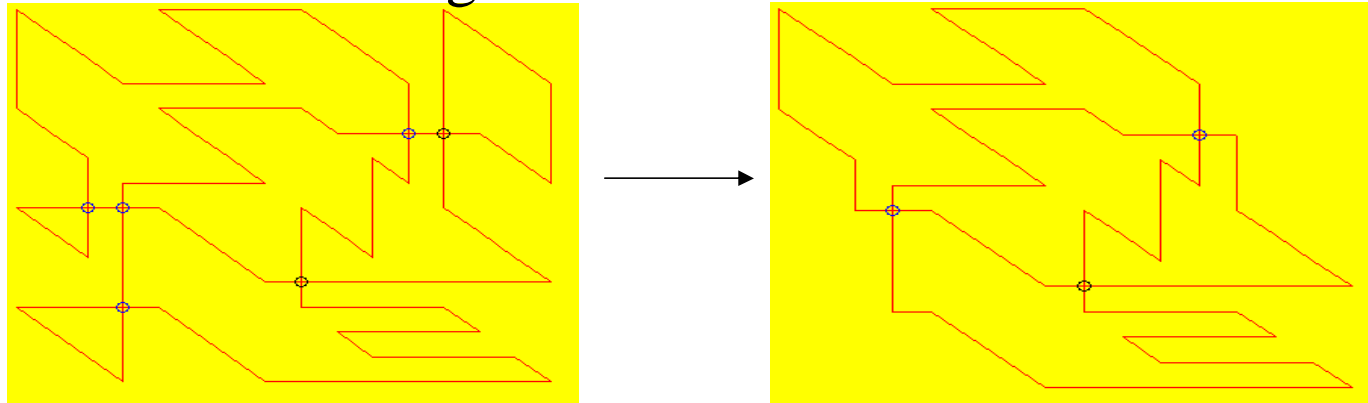
projection process



Projected nodes and links

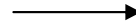
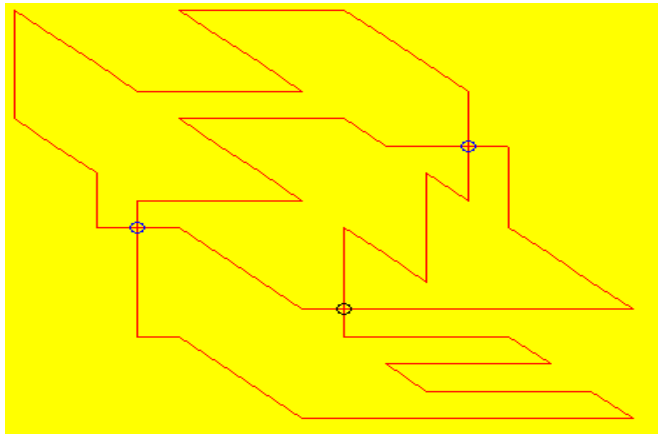
B. Preprocessing

Using Reidemeister moves

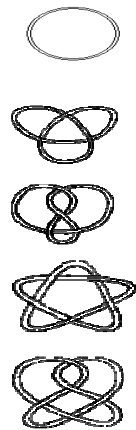
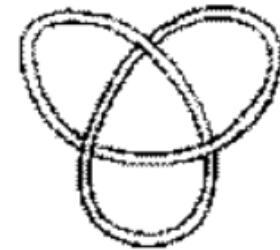


L	Average crossings before...	...and after reduction
4	20	0
6	114	4
8	383	40
10	969	187
12	2057	537
14	3896	1219

C. Knot Signature Computation



$$\begin{aligned}
 |\Delta_T(-1)| &= 3 && \text{(Alexander)} \\
 v_2(T) &= 1 && \text{(Vassiliev 2)} \\
 |v_3(T)| &= 1 && \text{(Vassiliev 3)}
 \end{aligned}$$



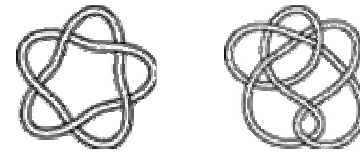
Knot	$ \Delta(-1) $	v_2	$ v_3 $	Chiral?
Trivial	1	0	0	NO
3-1	3	1	1	YES
4-1	5	-1	0	NO
5-1	5	3	5	YES
5-2	7	2	3	YES

Caveat!

Knot invariants cannot unambiguously classify a knot.

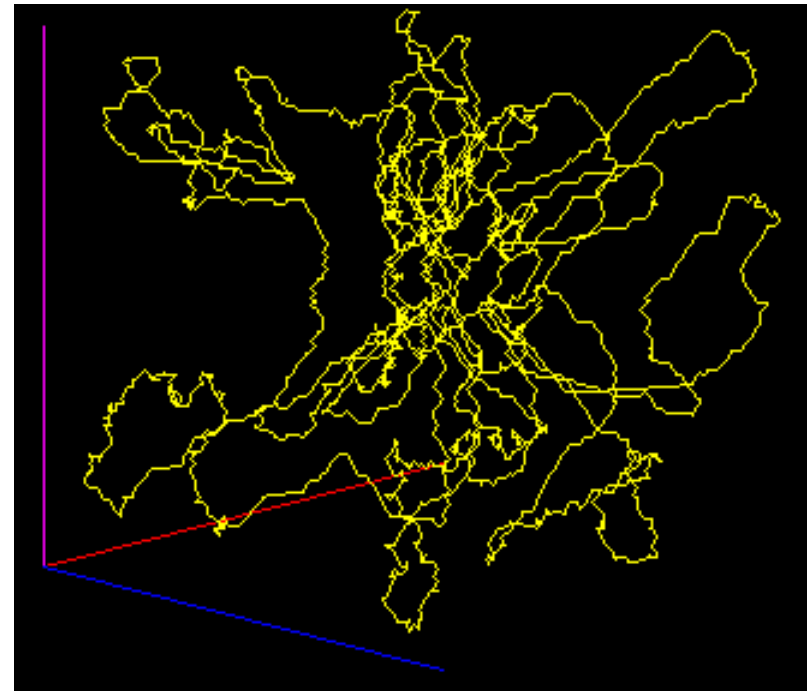
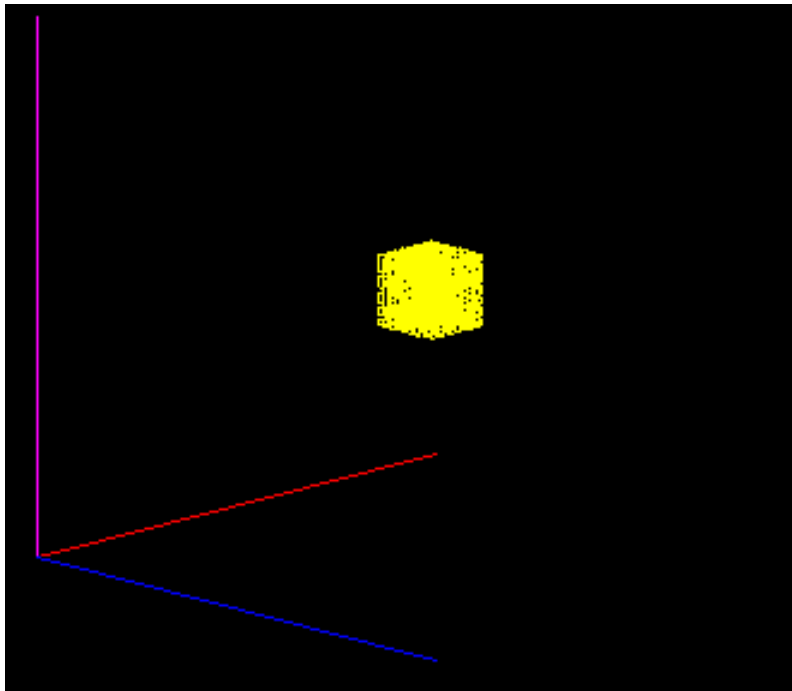
However

- knot invariants of the trivial knot and the first four knots are distinct from those of other prime knots with 10 crossings or fewer (249 knots in all), with one exception (5-1 and 10-132):



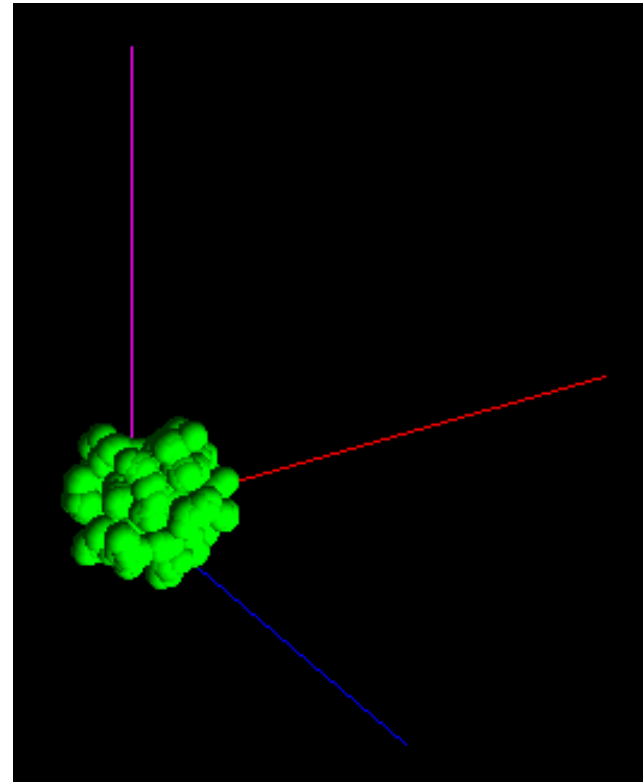
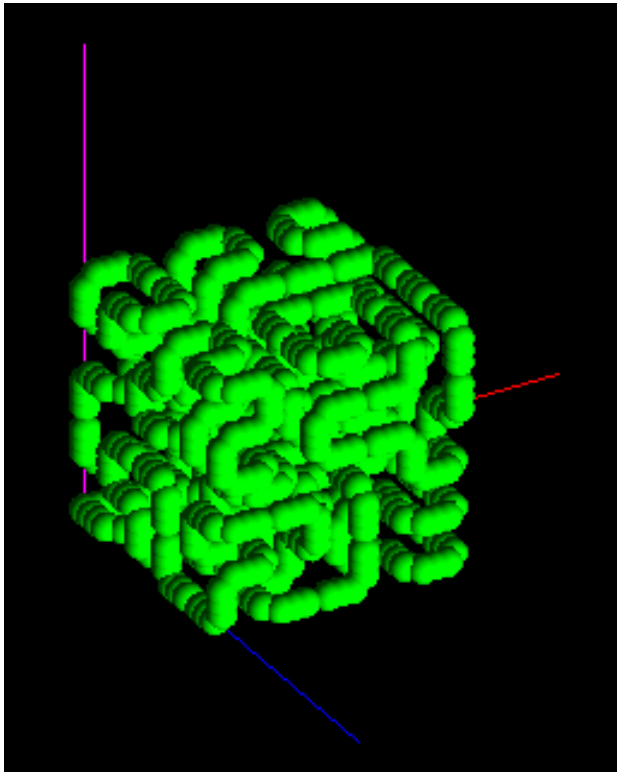
- Reidemeister moves and knot inflation can considerably reduce the number of possibilities.

Knot Inflation



Monte Carlo

Knot Tightening

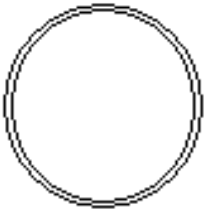


Shrink-On-No-Overlaps (SONO) method of Piotr Pieranski.
Scale all coordinates $s < 1$, keep bead radius fixed.

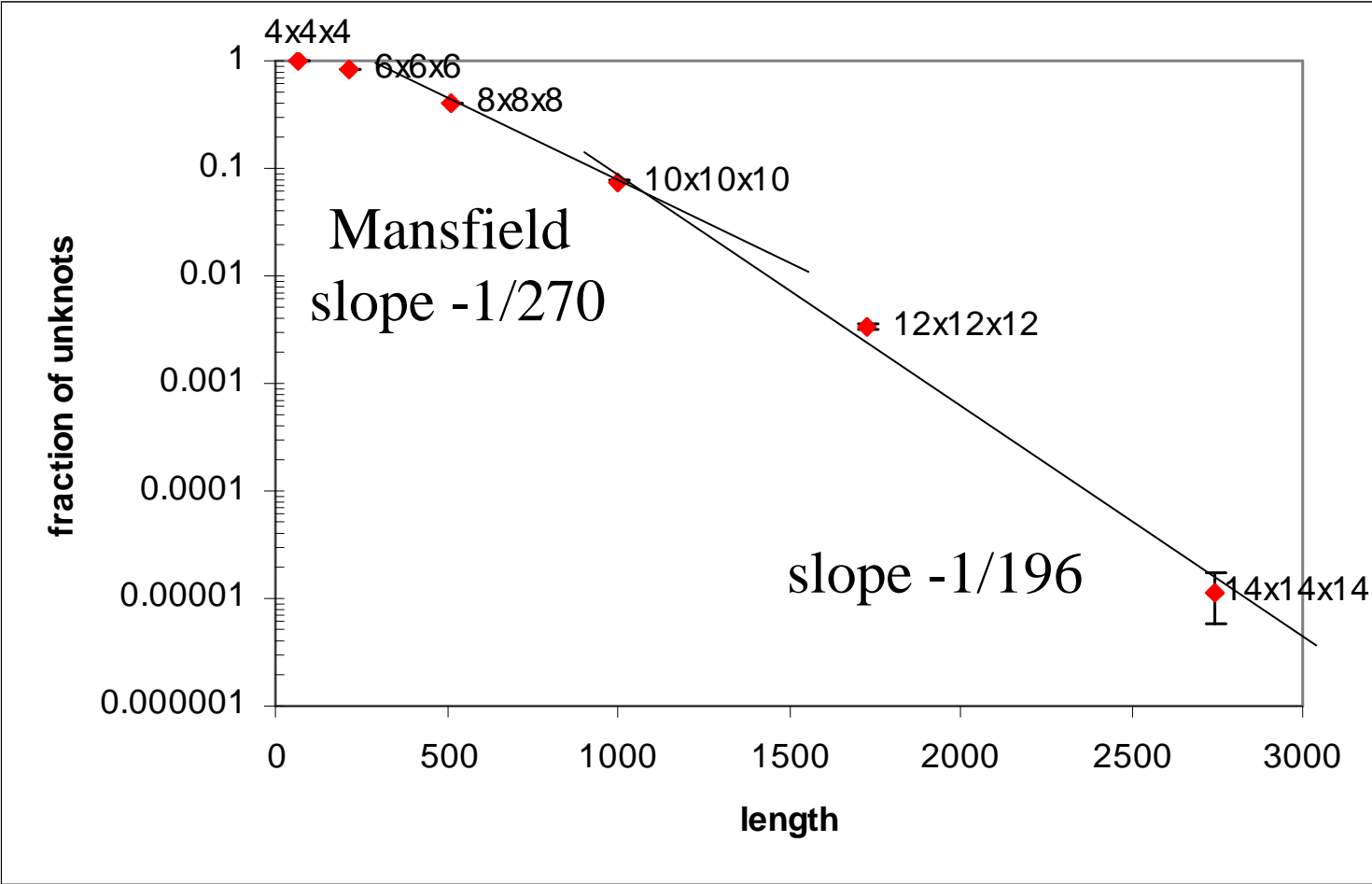
Results

Knotting Probabilities for Compact Lattice Loops

Chance of getting an unknot

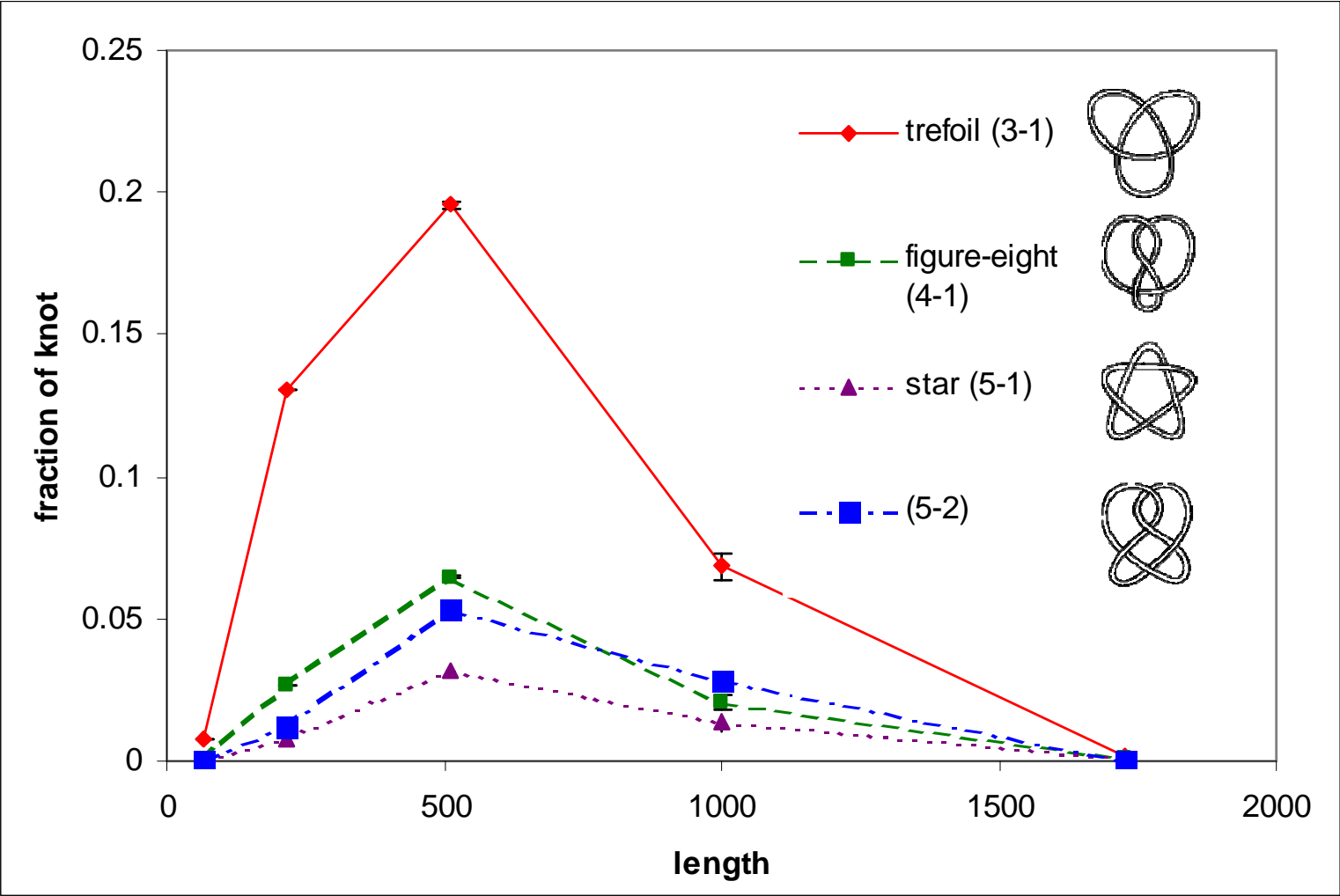


for several cube sizes:



$$P_0(N) \approx e^{-N/N_0} \quad N_0 \sim 200$$

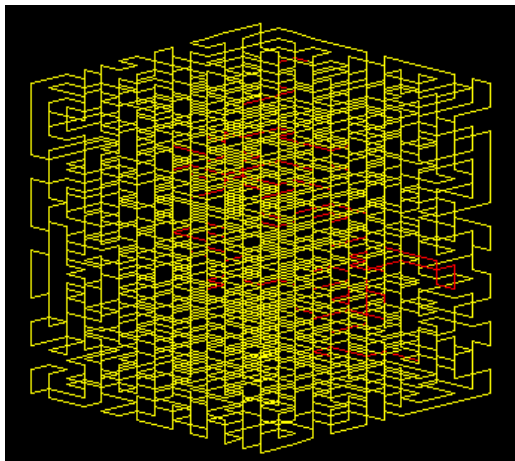
Chance of getting the first few simple knots for different cube sizes:



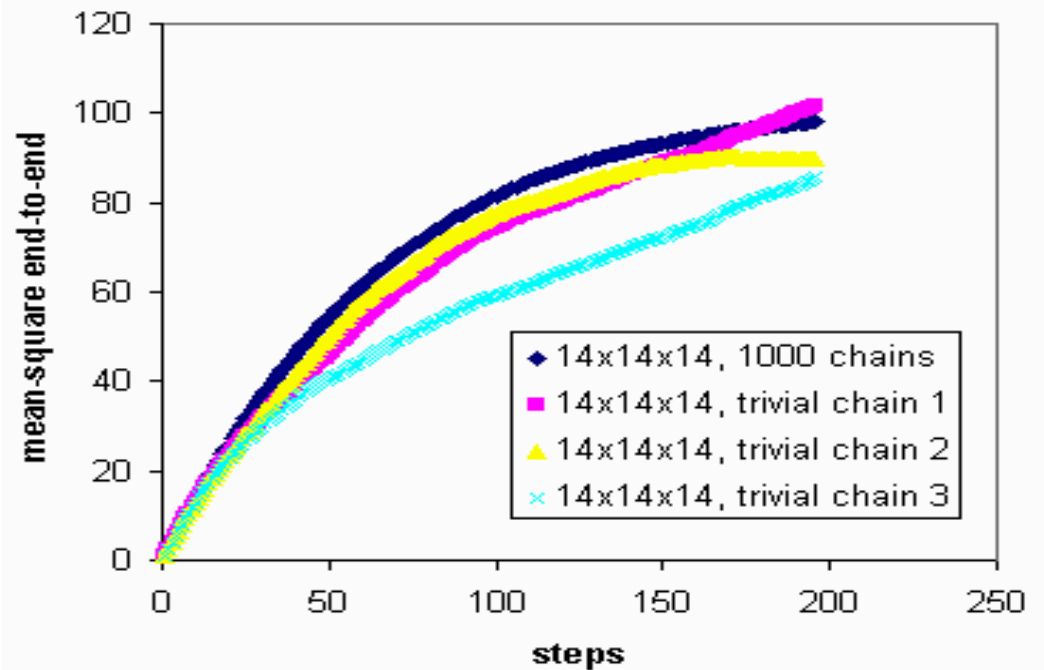
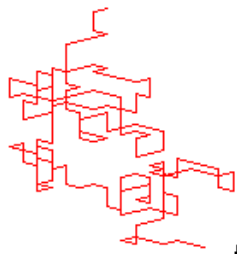
Subchain statistics

14x14x14 Compact Lattice Loop

Average size of subchain (mean-square end-to-end) versus length of subchain



Sub-chain:
(fragment)

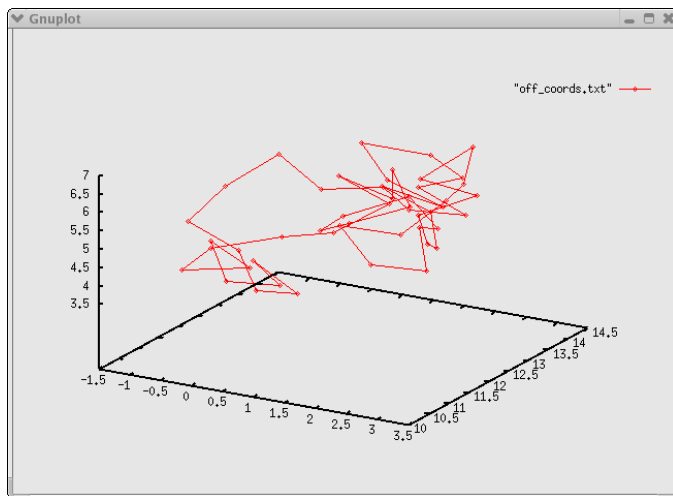


Fragments of trivial knots are more *crumpled* compared to fragments of all knots.

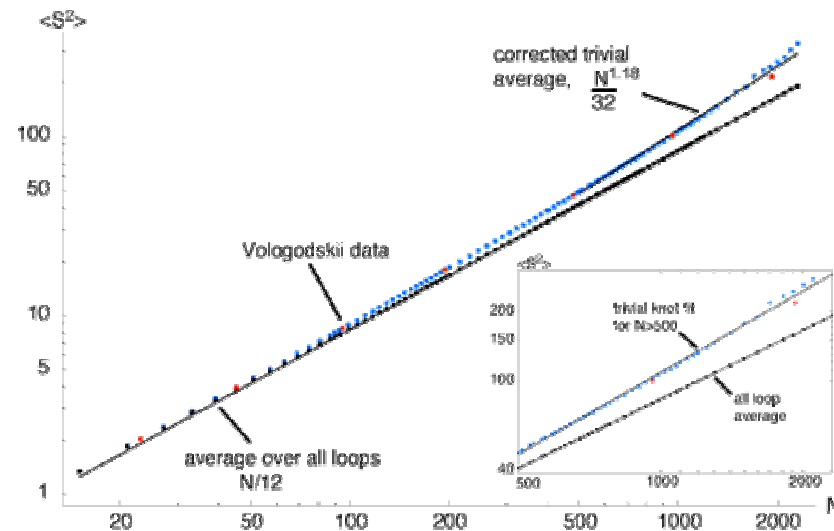
Noncompact, Unrestricted Loop

Average gyration radius (squared) versus length

Closed random walk
with fixed step length



Results: Gyration Radius Average



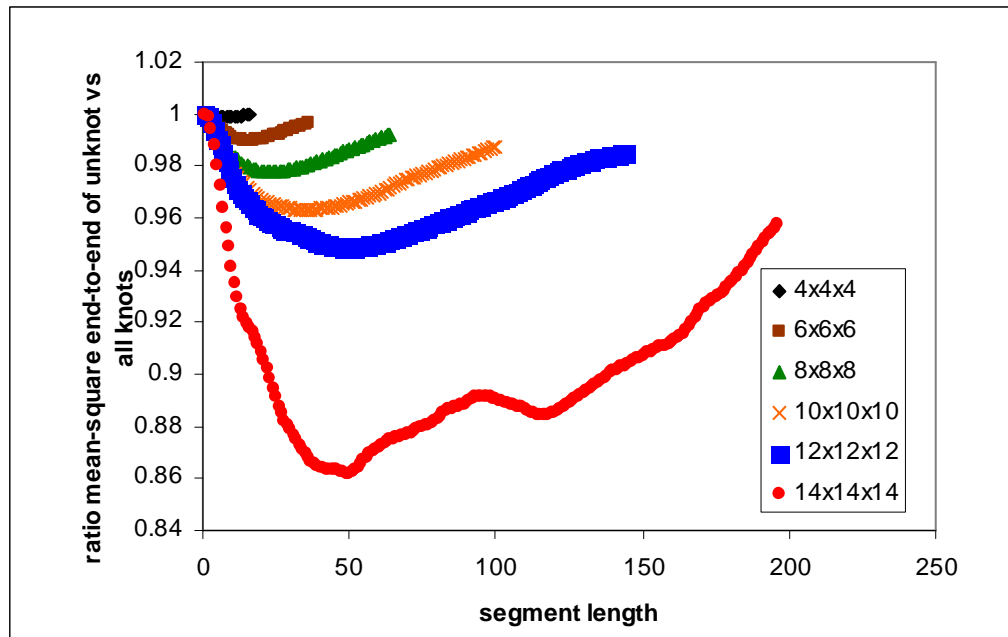
(N. Moore)

Trivial knots *swell* compared to all knots for noncompact chains.
This *topologically-driven* swelling is the same as that driven by self-avoidance (Flory exponent $3/5$ versus gaussian exponent $1/2$).

$$R \approx N^{\frac{3}{5}}$$

Compact Lattice Loops

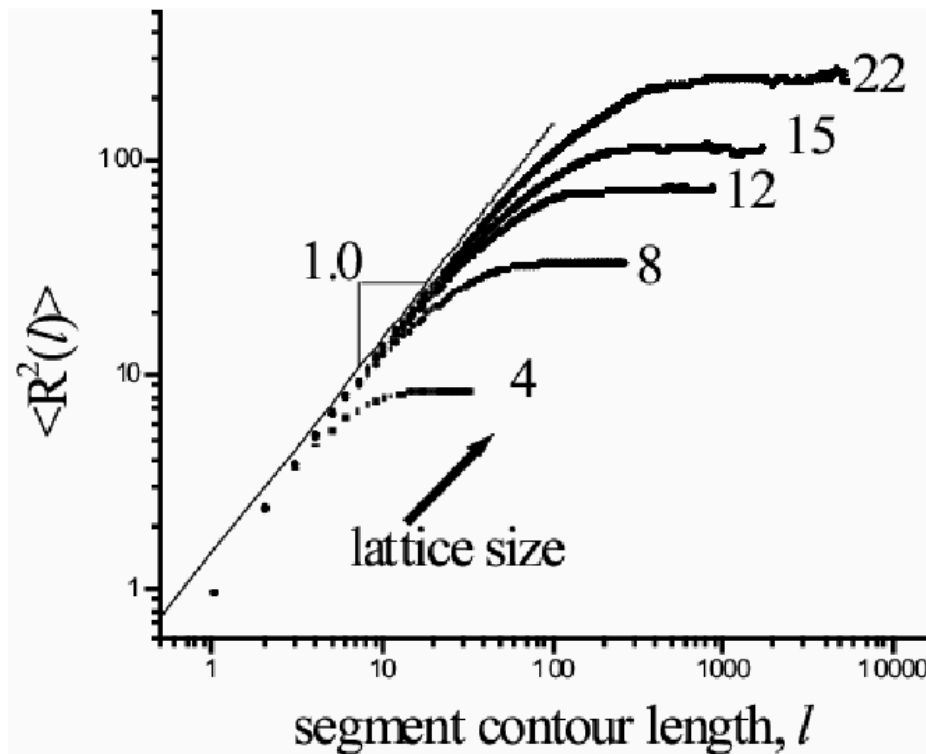
Ratios of average sub-chain sizes, trivial/all knots



Fragments of trivial knots are consistently more compact compared to fragments of all knots.

Compact Lattice Loops

General scaling of subchains (mean-square end-to-end) versus length



(A. Borovinskiy)

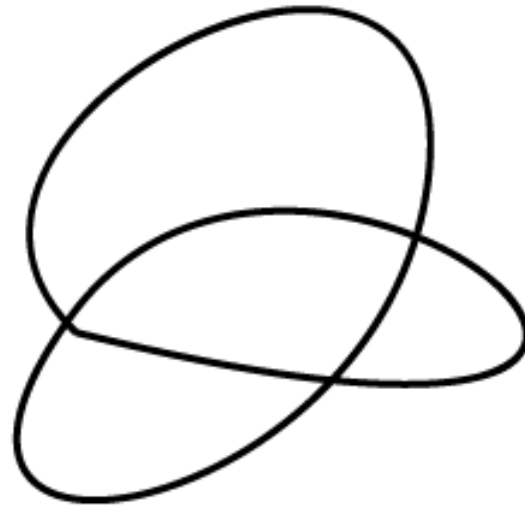
Over all knots: $R \approx t^{\frac{1}{2}}$

i.e. Gaussian;
Flory's result for chains in
a polymer melt.

Trivial knots: $R \approx t^{\frac{1}{3}}$?

Knot (De)Localization

Localized or delocalized?



What have been shown computationally...

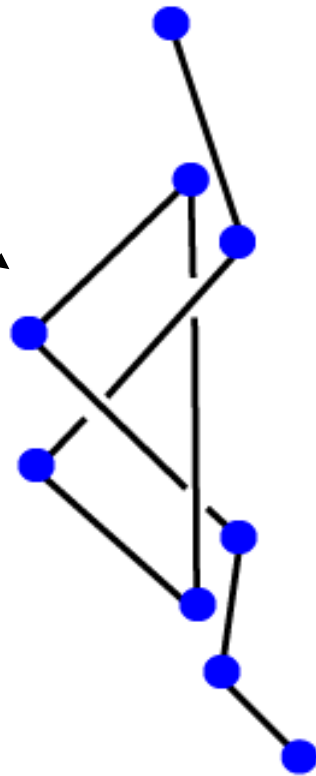
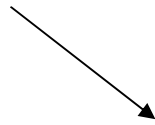
	Random circular chains	Flat knots (Polymer on sticky surface)
Noncompact (swollen)	Localized*	Localized**
Compact (collapsed)	?	Delocalized**

*Katritch, Olson, Vologodskii, Dubochet, Stasiak (2000). Preferred size of 'core' of trefoil knot is 7 segments.

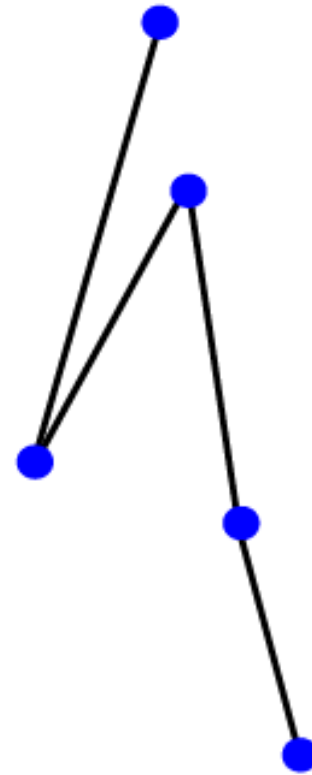
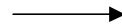
**Orlandini, Stella, Vanderzande (2003). Localization to delocalization transition below a θ -point temperature.

Knot Renormalization

Localized
trefoil

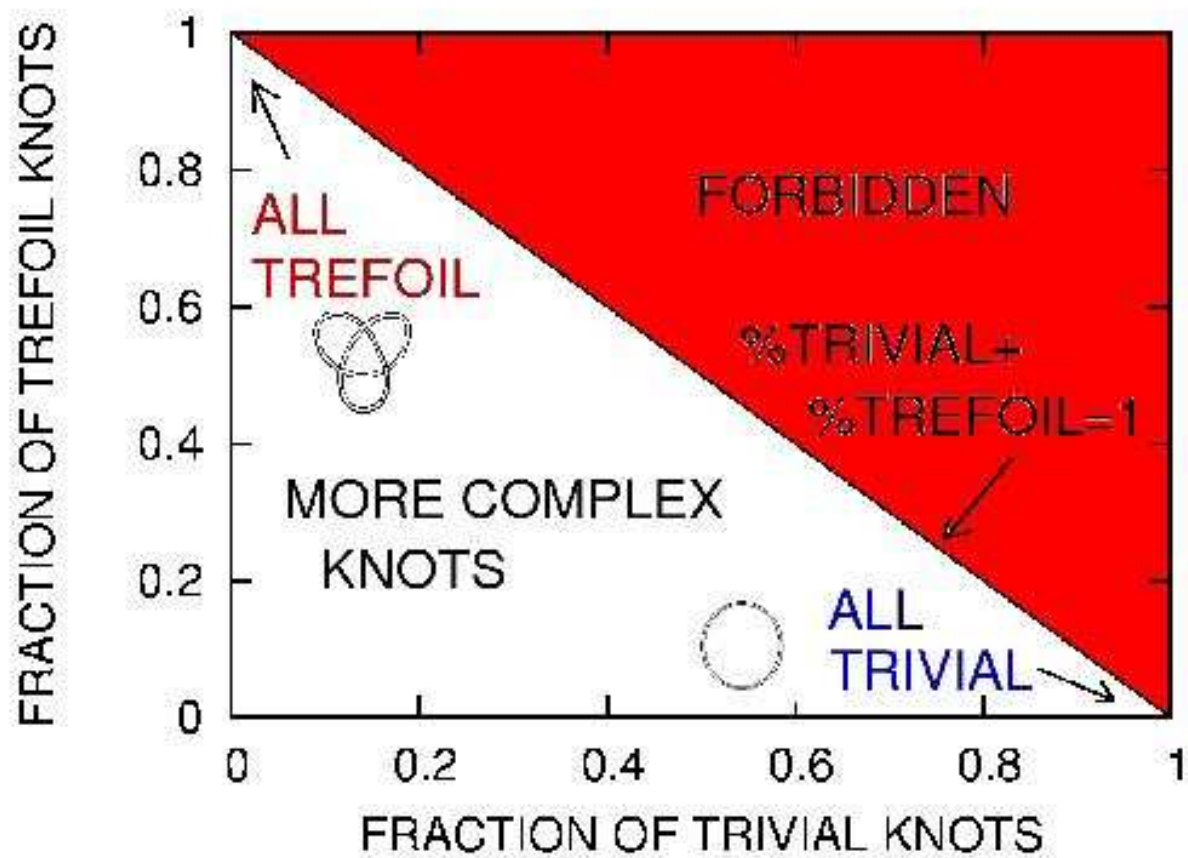


$g=1$



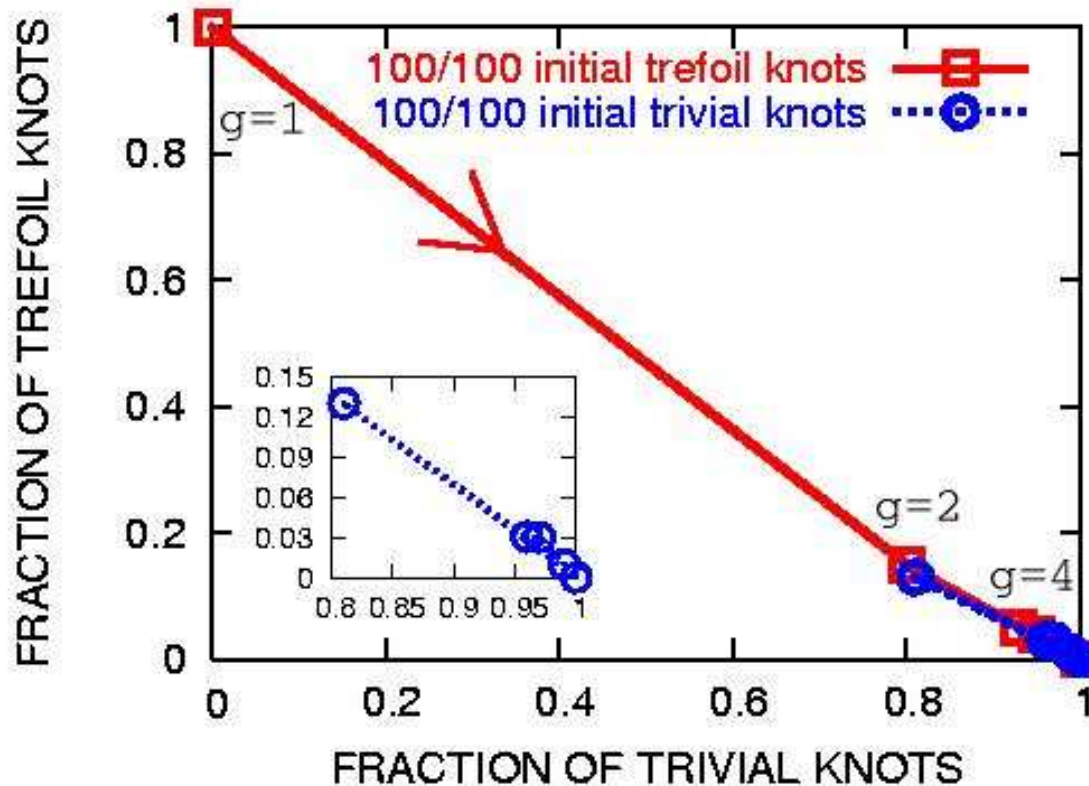
$g=2$

Renormalization trajectory space



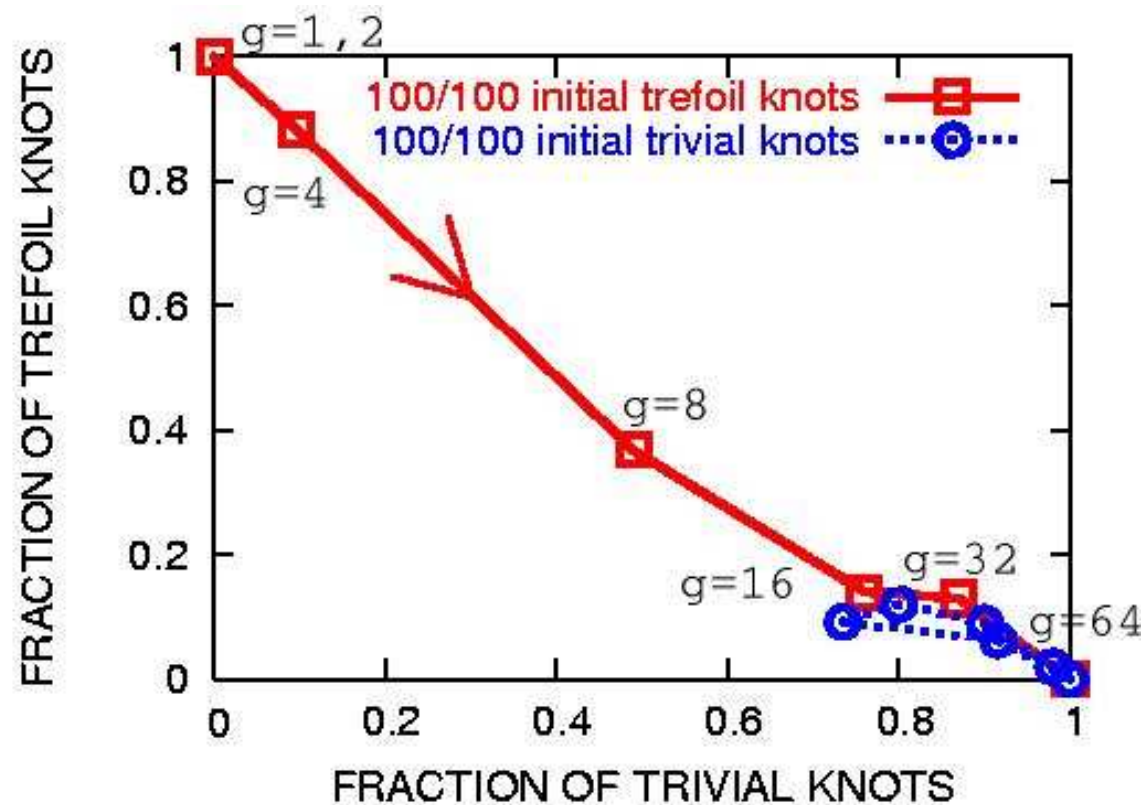
Renormalization trajectory

Initial state: Noncompact loop, $N=384$



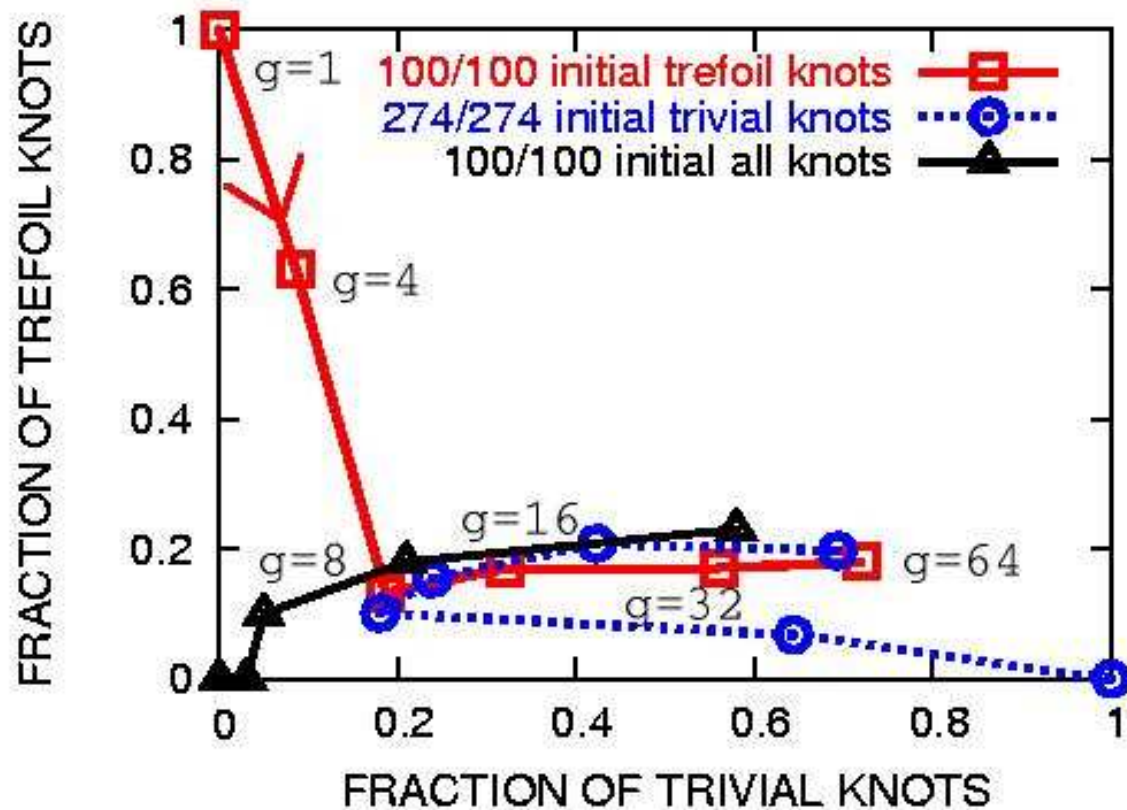
Renormalization trajectory

Initial state: 8x8x8 compact lattice loop



Renormalization trajectory

Initial state: 12x12x12 compact lattice loop



Knots in Proteins

Previous work...

1. M.L. Mansfield (1994):

Approx. 400 proteins, with random bridging of terminals, using Alexander polynomial. Found at most 3 knots.

2. W.R. Taylor (2000)

3440 proteins, fixing the terminals and smoothing (shrinking) the segments in between. Found 6 trefoils and 2 figure-eights.

3. K. Millet, A. Dobay, A. Stasiak (2005)

(Not about proteins) A study of linear random knots and their scaling behaviour.

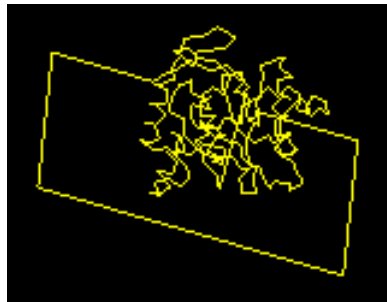
Steps

1. Obtain protein structural information (.pdb files) from the Protein Data Bank. 4716 id's of representative protein chains obtained from the Parallel Protein Information Analysis (PAPIA) system's Representative Protein Chains from PDB (PDB-REPRDB).

2. Extract coordinates of protein backbone



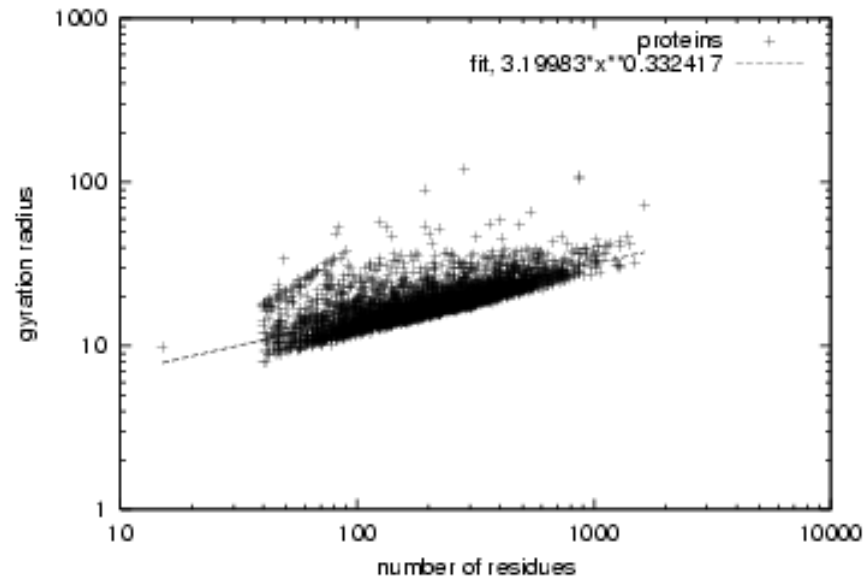
3. Close the knot



(3 ways)

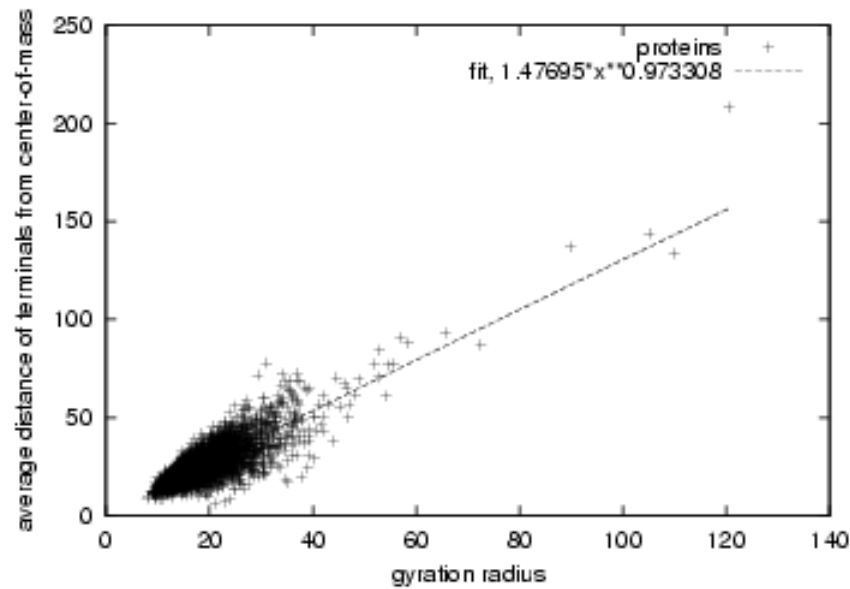
4. Calculate knot invariants/signatures

Protein gyration radius versus length



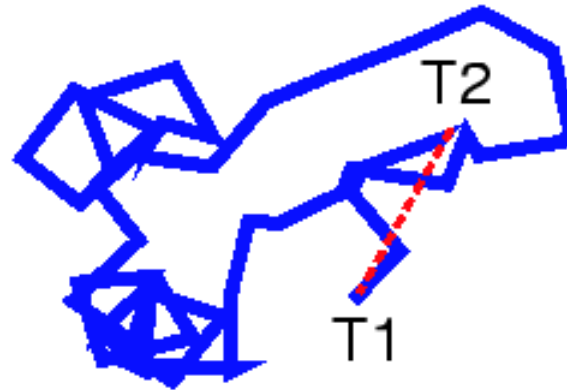
$$R = aN^{\frac{1}{3}}$$

CM-to-terminals distance versus gyration radius



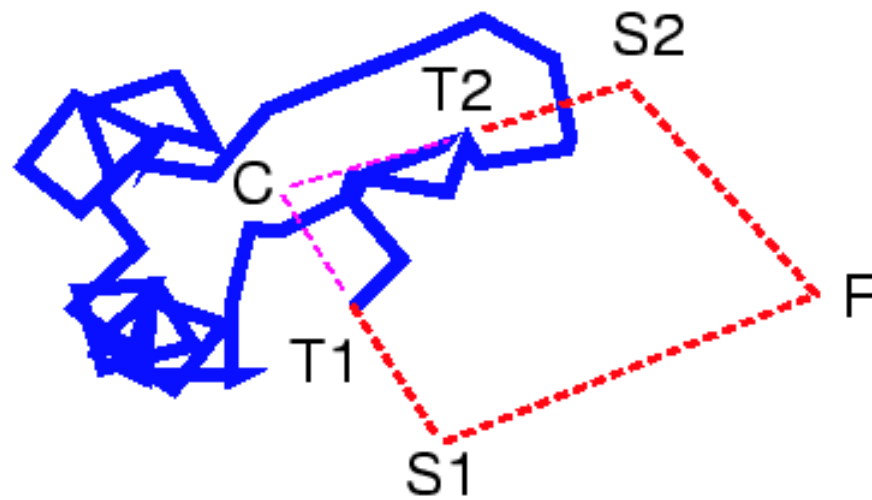
$$\langle C - T \rangle \approx 1.5R$$

DIRECT closure method



T1, T2 – protein terminals

CM-AYG closure method

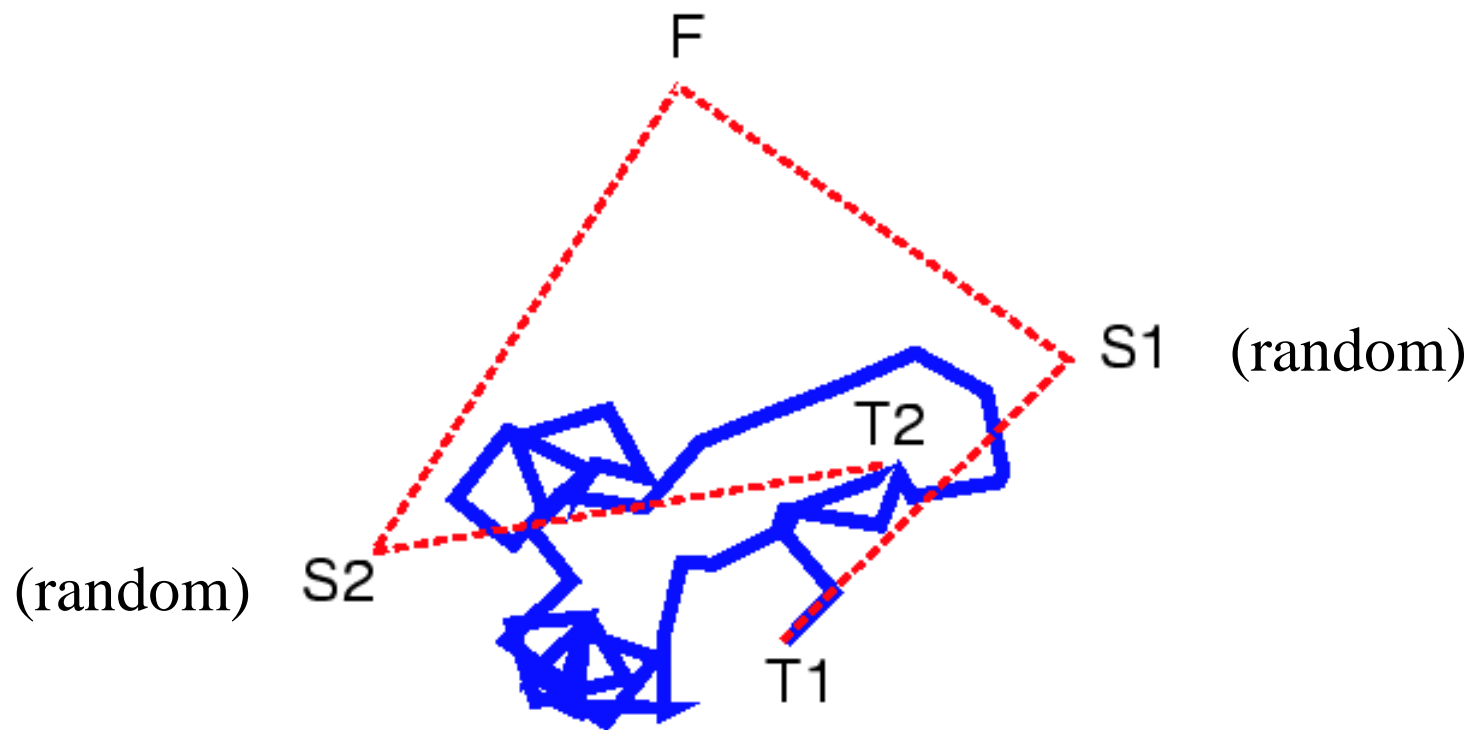


C – center of mass

S1, S2 - located on surface of sphere surrounding the protein

F- point at some large distance away from C

RANDOM2 closure method

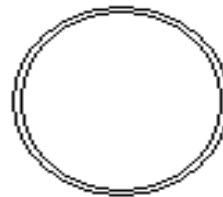
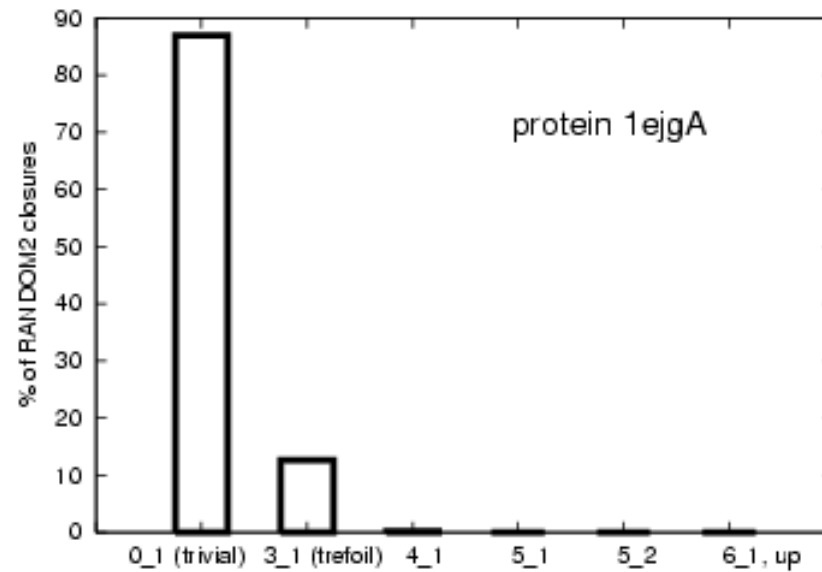


Study statistics of knot closures after generating 1000 pairs of points S1 and S2. Determine the dominant knot-type.

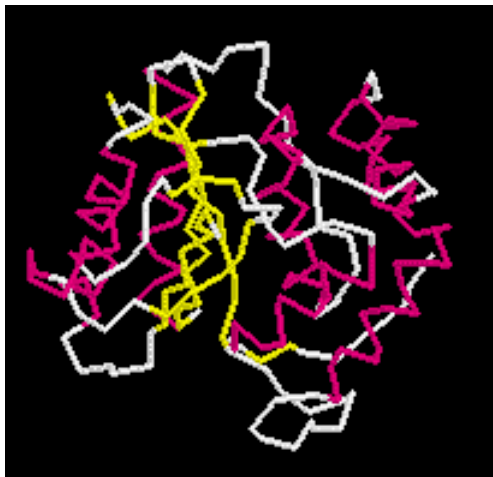
Knot probabilities in RANDOM2 closures for protein 1ejg chain A



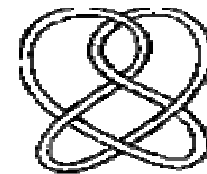
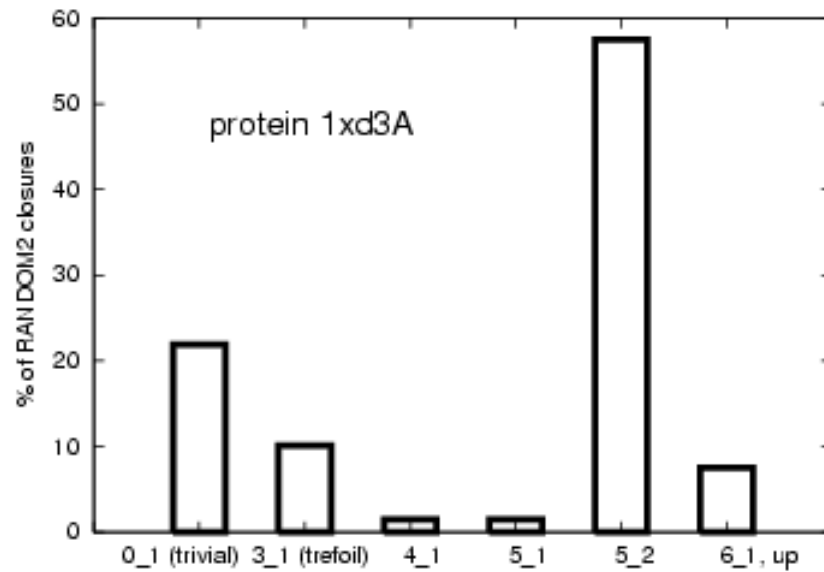
N=46



Knot probabilities in RANDOM2 closures for protein 1xd3 chain A



N=229

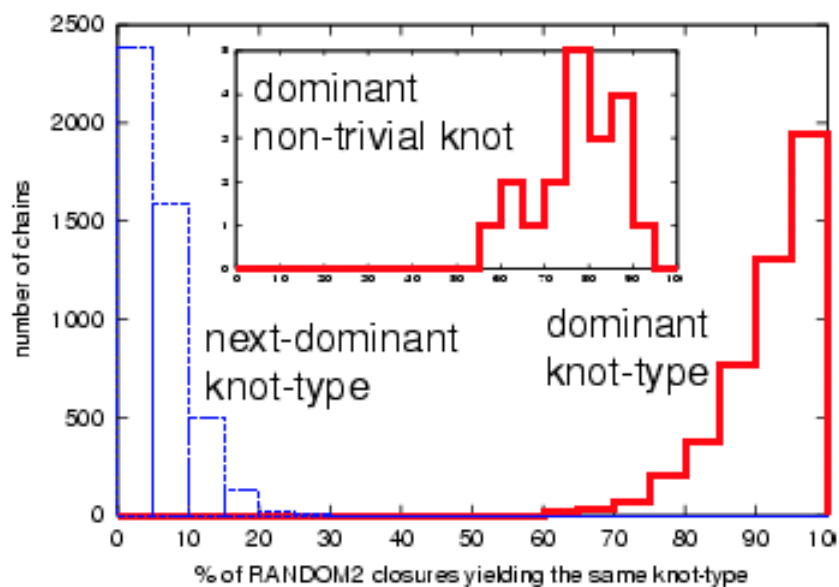


Knot counts in the three closure methods

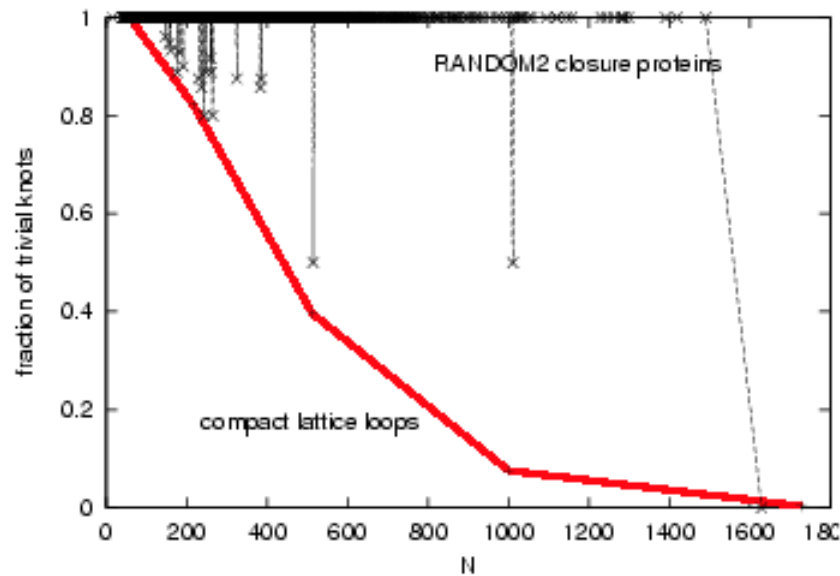
Closure method	Knot-type					
	0_1	3_1	4_1	5_1	5_2	$6_{1,up}$
RANDOM2	4697	15	3	0	1	0
CM-AYG	4692	20	3	0	1	0
DIRECT	4516	164	9	9	3	15

- RANDOM2 and CM-AYG methods gave the same predictions for 4711 chains (out of 4716).
- RANDOM2 and DIRECT methods gave the same predictions for 4528 chains (out of 4716).

Distribution of the % of RANDOM2 closures giving the dominant knot-type



Unknotting probabilities versus length for proteins and for compact lattice loops



Total of 19 non-trivial knots in the RANDOM2 method.

Knots in proteins occur much less often than in compact lattice loops.

Summary of Results

- Unknotting probability drops exponentially with chain length.

$$P_{trivial} \approx e^{-N/196}$$

- For compact conformations, subchains of trivial knots are consistently smaller than subchains of non-trivial knots. For noncompact conformations, the opposite is observed. *The fragments seem to be 'aware' of the knottedness of the whole thing.*

(AYG)

- Knots in proteins are rare.

Unresolved issues...

- Are knots in compact loops delocalized? To what degree?
- Theoretical treatment of the scaling of subchains in compact loops with trivial knots.
- Theoretical prediction for the characteristic length of knotting N_0 . $P_0(N) \approx e^{-N/N_0}$

Acknowledgments

A. Yu. Grosberg.

A. Borovinskiy, N. Moore.

P. Pieranski and associates for
SONO animation.

DDF support, UMN Graduate School
Knot Mathematicians.

Biologists, Chemists and other researchers for
making protein structures available.