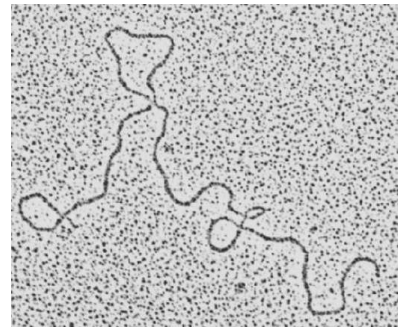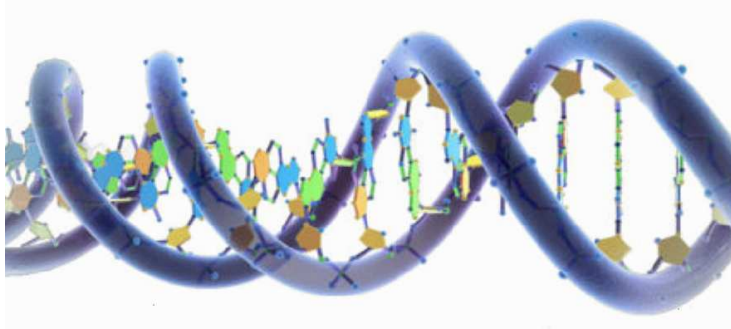# Predicting local geometric properties of DNA from hydrodynamic diffusion data
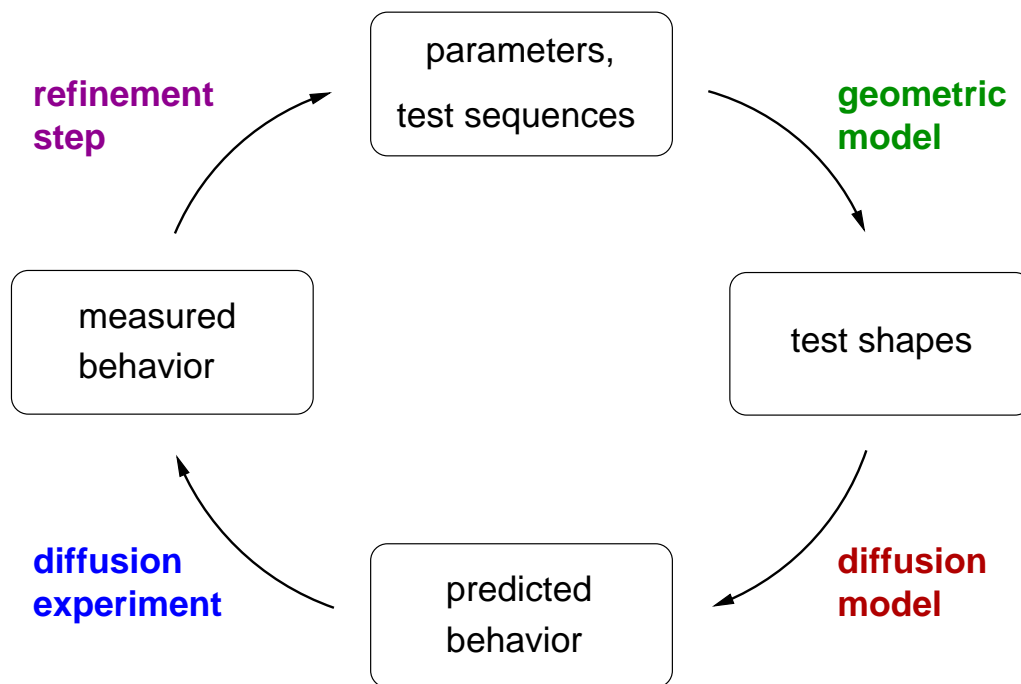
O. Gonzalez,   J. Li
Department of Mathematics
University of Texas, Austin

Introduction

Goal: To obtain improved estimates of the material parameters that describe sequence-dependent shape of DNA.

Approach: Refine parameters through hydrodynamic diffusion modeling and experiments.
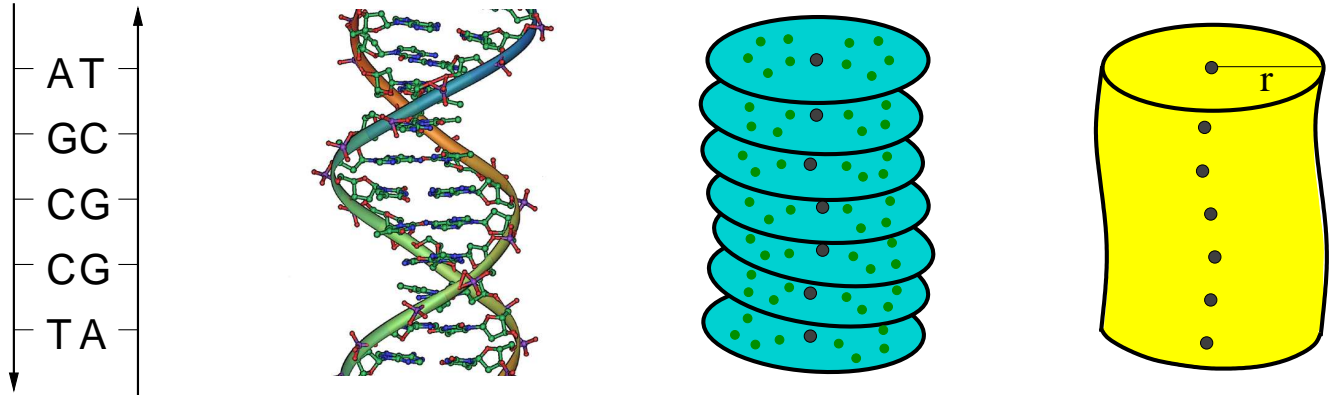
Outline

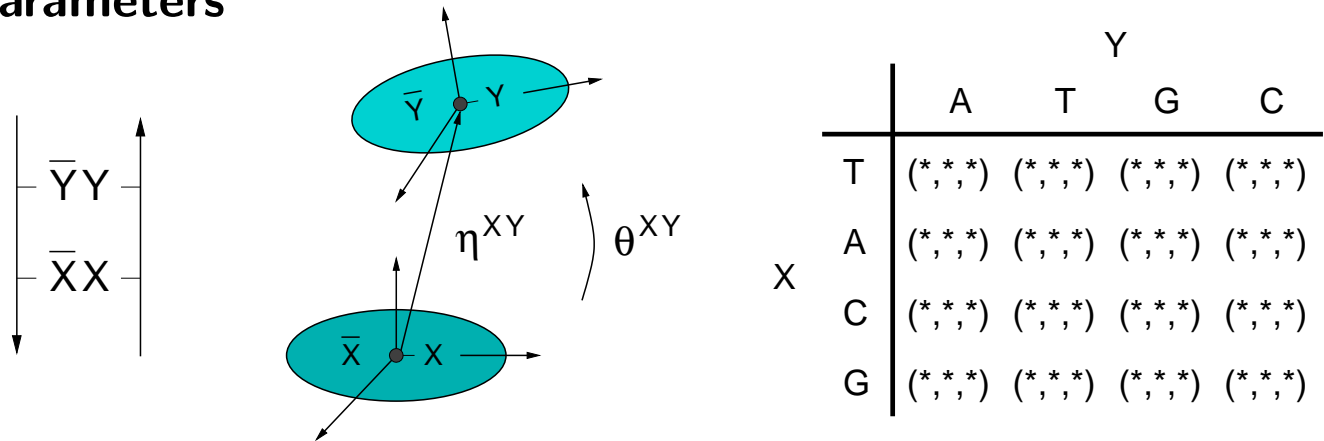**Geometric model**

Diffusion model

Application

Results

Geometric model of DNA

Definition

A T
G C
C G
C G
T A

Parameters

$\overline{Y}Y$

$\overline{X}X$

$\overline{Y}$ — Y

η^{XY} θ^{XY}

$\overline{X}$ — X

		Y			
		A	T	G	C
X	T	(*,*,*)	(*,*,*)	(*,*,*)	(*,*,*)
	A	(*,*,*)	(*,*,*)	(*,*,*)	(*,*,*)
	C	(*,*,*)	(*,*,*)	(*,*,*)	(*,*,*)
	G	(*,*,*)	(*,*,*)	(*,*,*)	(*,*,*)

η^{XY} = relative displacement for dimer step XY.

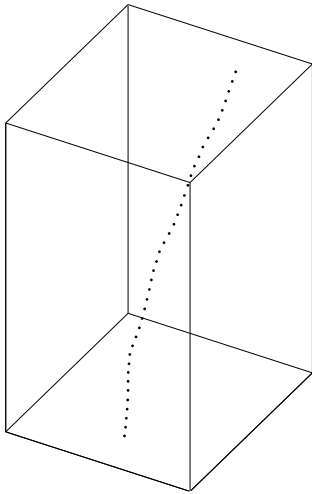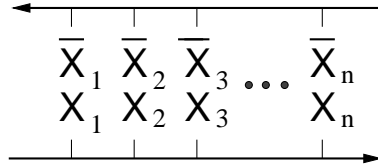θ^{XY} = relative rotation for dimer step XY.

r = radius of tubular surface.

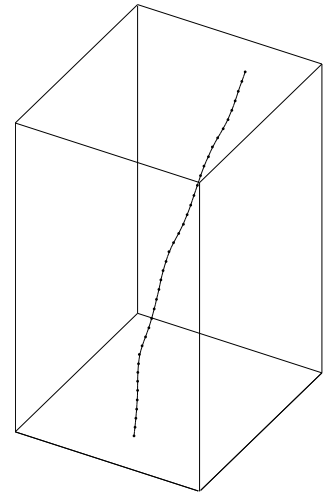$$\eta = \{\eta^{XY}\} \in \boldsymbol{R}^{26}, \quad \theta = \{\theta^{XY}\} \in \boldsymbol{R}^{26}, \quad r \in \boldsymbol{R}.$$

Geometric model of DNA

Construction

$$S = \mathsf{X}_1\mathsf{X}_2\cdots\mathsf{X}_n$$

$$\begin{array}{ccccc} \overline{\mathsf{X}}_1 & \overline{\mathsf{X}}_2 & \overline{\mathsf{X}}_3 & \cdots & \overline{\mathsf{X}}_n \\ \mathsf{X}_1 & \mathsf{X}_2 & \mathsf{X}_3 & & \mathsf{X}_n \end{array}$$
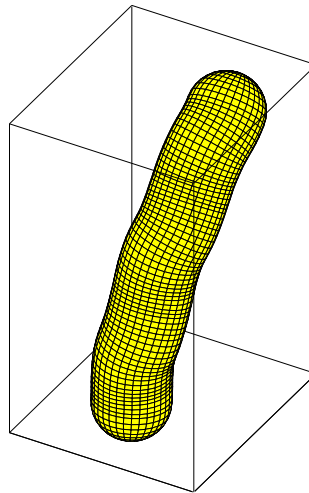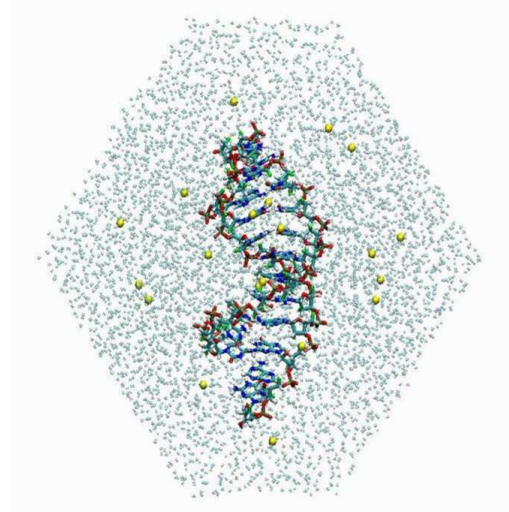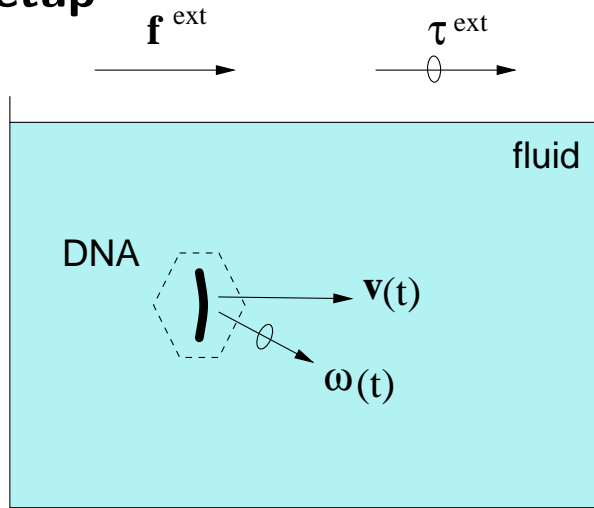
Reference points $q_a(S, \eta, \theta)$.

Axial curve $\gamma(S, \eta, \theta)$.

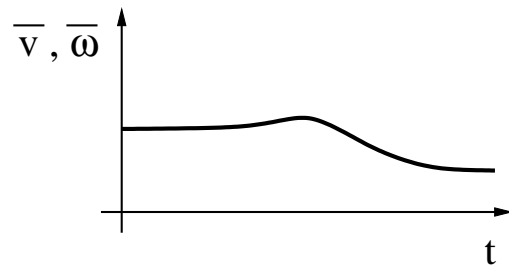Tubular surface $\Gamma(S, \eta, \theta, r)$.
(assumed rigid)

Diffusion model:
single molecule drift dynamics

Setup

Drift velocities

$$\overline{v}(t), \overline{\omega}(t) = \frac{1}{T} \int_t^{t+T} v(s), \omega(s) \ ds. \qquad T = \text{window size.}$$
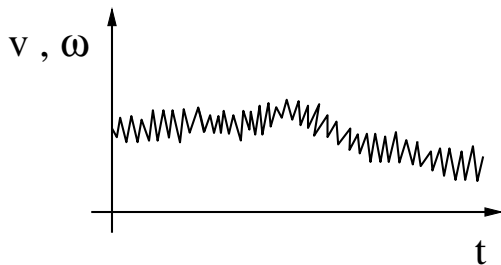
Governing equation

$$\begin{bmatrix} \overline{v} \\ \overline{\omega} \end{bmatrix} = M \begin{bmatrix} f \\ \tau \end{bmatrix}^{\text{ext}}.$$

$M \in \boldsymbol{R}^{6 \times 6}$ Stokes mobility matrix
determined by shape of DNA molecule.

Diffusion model:
single molecule mobility matrix

Stokes equations
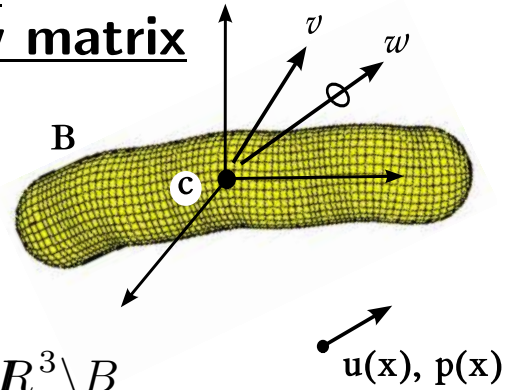
$$
\begin{aligned}
\mu \Delta \overline{u} &= \nabla \overline{p} && \text{in } \boldsymbol{R}^3 \backslash B \\
\nabla \cdot \overline{u} &= 0 && \text{in } \boldsymbol{R}^3 \backslash B \\
\overline{u} &= \overline{U}[\overline{v}, \overline{\omega}] && \text{on } \partial B \\
\overline{u}, \overline{p} &\to 0 && \text{as } |x| \to \infty.
\end{aligned}
$$

BCs, loads

$$
\overline{U}[\overline{v}, \overline{\omega}] = \overline{v} + \overline{\omega} \times (x - c)
$$

$$
\overline{\sigma}_{ij}[\overline{u}, \overline{p}] = -\overline{p}\delta_{ij} + \mu \left(\frac{\partial \overline{u}_i}{\partial x_j} + \frac{\partial \overline{u}_j}{\partial x_i} \right)
$$

$$
\overline{f}[\overline{\sigma}] = \int_{\partial B} \overline{\sigma} n \, dA, \qquad \overline{\tau}[\overline{\sigma}] = \int_{\partial B} (x - c) \times \overline{\sigma} n \, dA.
$$

Hydrodynamic matrices

$$
\begin{bmatrix} \overline{f} \\ \overline{\tau} \end{bmatrix} = - \underbrace{\begin{bmatrix} L_1 & L_3 \\ L_2 & L_4 \end{bmatrix}}_{L} \begin{bmatrix} \overline{v} \\ \overline{\omega} \end{bmatrix}, \qquad \begin{bmatrix} \overline{v} \\ \overline{\omega} \end{bmatrix} = - \underbrace{\begin{bmatrix} M_1 & M_3 \\ M_2 & M_4 \end{bmatrix}}_{M} \begin{bmatrix} \overline{f} \\ \overline{\tau} \end{bmatrix}.
$$

Diffusion model: collective drift dynamics

Setup

Physical domain

$\mathbf{f}^{\text{ext}}$

τ^{ext}

Dom

Configuration space

(r,Q)

Dom x SO$_3$

Local coord space

(q,η)

Ω x A
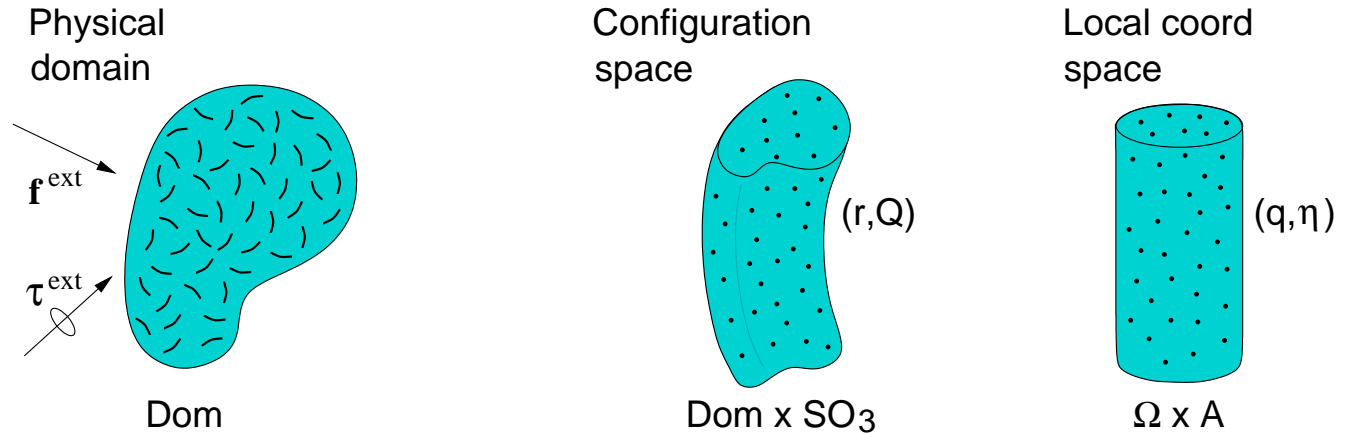
$\sigma =$ DNA concentration in config space.

$h^{\text{ext}} = (f, \tau)^{\text{ext}} =$ external loads. $M =$ Stokes mobility matrix.

Governing equation

$$\sigma_t + g^{-1}\nabla \cdot (gJ) = 0 \quad \text{in} \quad \Omega \times A, \qquad t > 0$$
$$J \cdot n = 0 \quad \text{on} \quad (\partial \Omega) \times A, \quad t \geq 0$$
$$J \cdot n \quad \text{periodic} \quad \text{on} \quad \Omega \times (\partial A), \quad t \geq 0$$
$$\sigma = \sigma_0 \quad \text{in} \quad \Omega \times A, \qquad t = 0.$$

$$J = -D\nabla\sigma + \sigma C h^{\text{ext}}, \quad D = \beta b M b^T, \quad C = bMc.$$
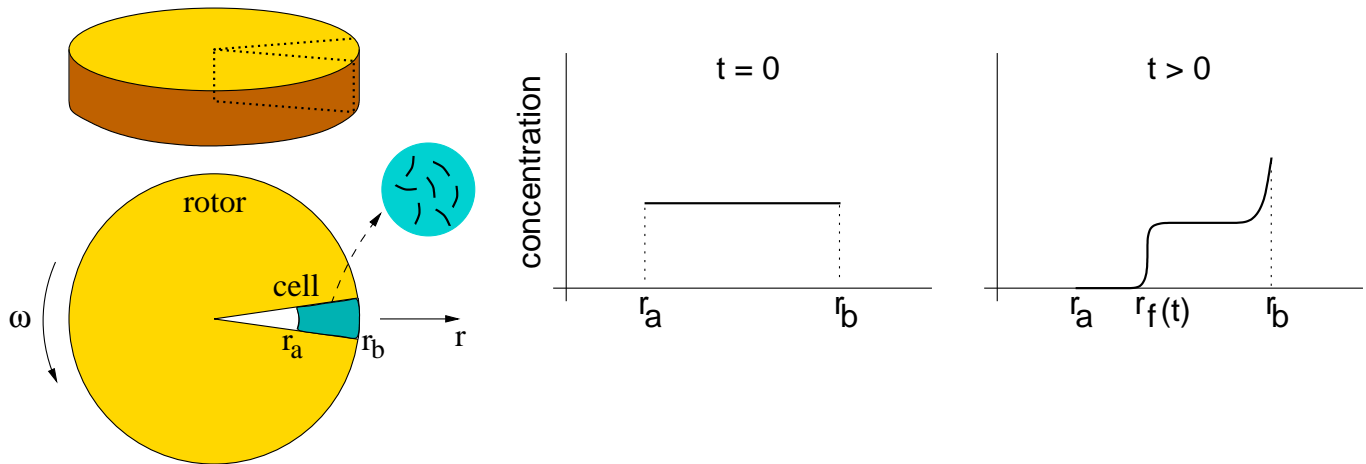
$\beta =$ Boltzmann factor. $g, b, c =$ geometric factors.

$D, C \in \mathbb{R}^{6 \times 6}$ local diffusion, convection matrices determined by shape of DNA molecule.

Diffusion model: example experiment

Ultracentrifuge

Governing eq (dimless).

$$\sigma_t = g^{-1}\nabla \cdot \left[gD(\tfrac{1}{\varepsilon})\nabla\sigma - g\sigma C(\tfrac{1}{\varepsilon})h^{\mathrm{ext}} \right] \quad \text{in} \quad \Omega \times A.$$
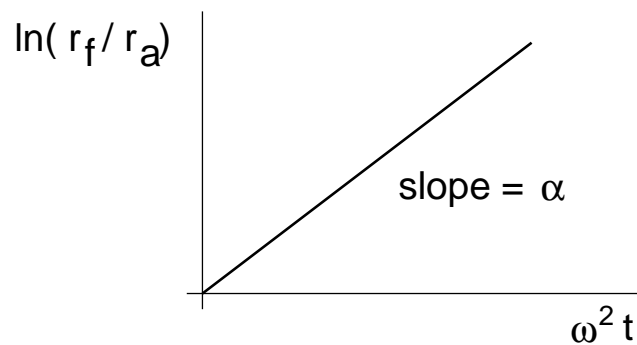
$$\varepsilon << 1.$$

Leading-order approx.

$$\frac{\partial\rho}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left[\kappa\, r\frac{\partial\rho}{\partial r} - \alpha\,\omega^2 r^2\rho \right] \quad \text{in} \quad [r_a, r_b].$$

$$\kappa = f(D), \qquad \alpha = \widehat{f}(C).$$

Result.

Diffusion model:
summary

- Different diffusion experiments can measure different components of $D, C \in \boldsymbol{R}^{6 \times 6}$, or equivalently, $M \in \boldsymbol{R}^{6 \times 6}$.

- A typical measurable quantity is the translational diffusion coefficient

$$D_t = \frac{\beta}{3} \operatorname{tr}(M_1) \quad \text{where} \quad M = \begin{bmatrix} M_1 & M_3 \\ M_2 & M_4 \end{bmatrix}.$$
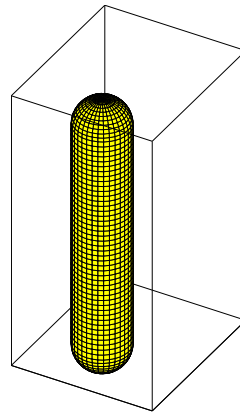
- M and D_t can be predicted by solving Stokes flow equations around given molecular geometry.
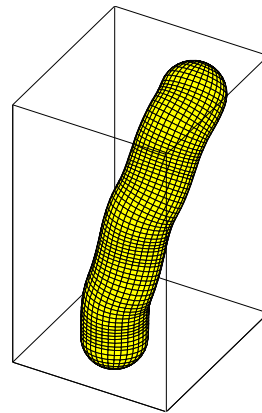
Application

Given: Measured values of D_t on sequences of different length.

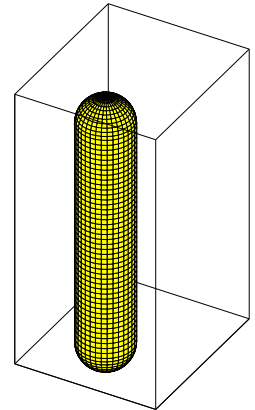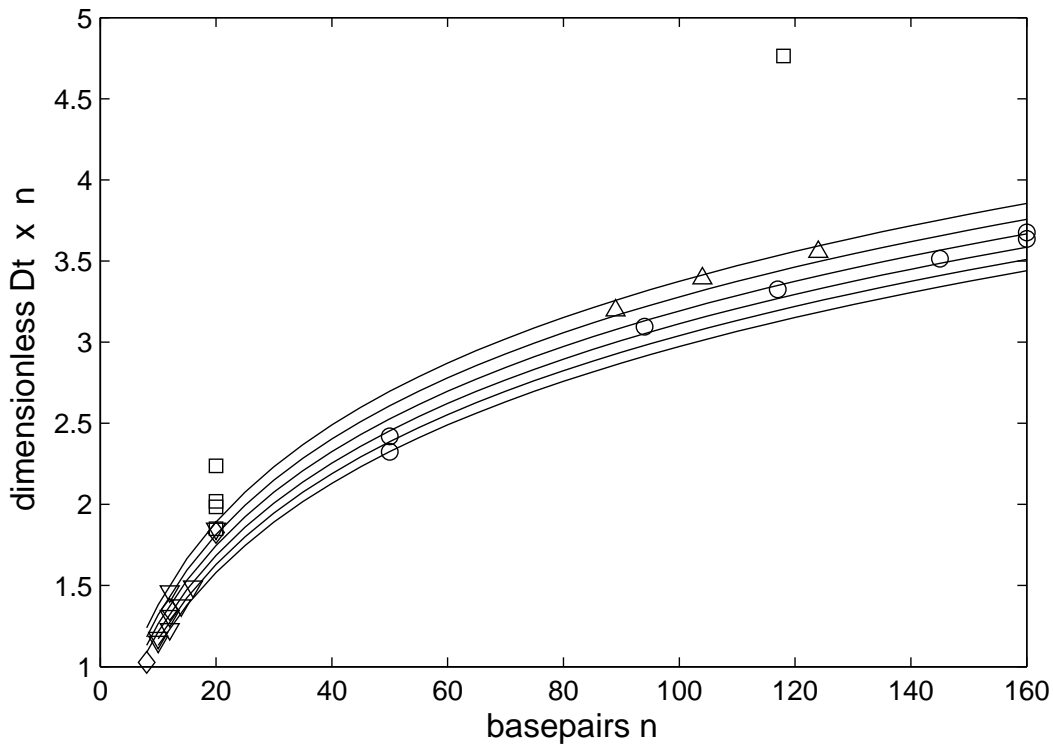**Find:** Radius r of (hydrated) DNA using two different geometrical models:

Idealized (straight) model.
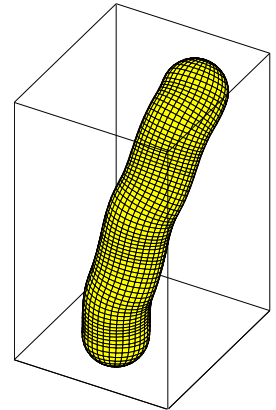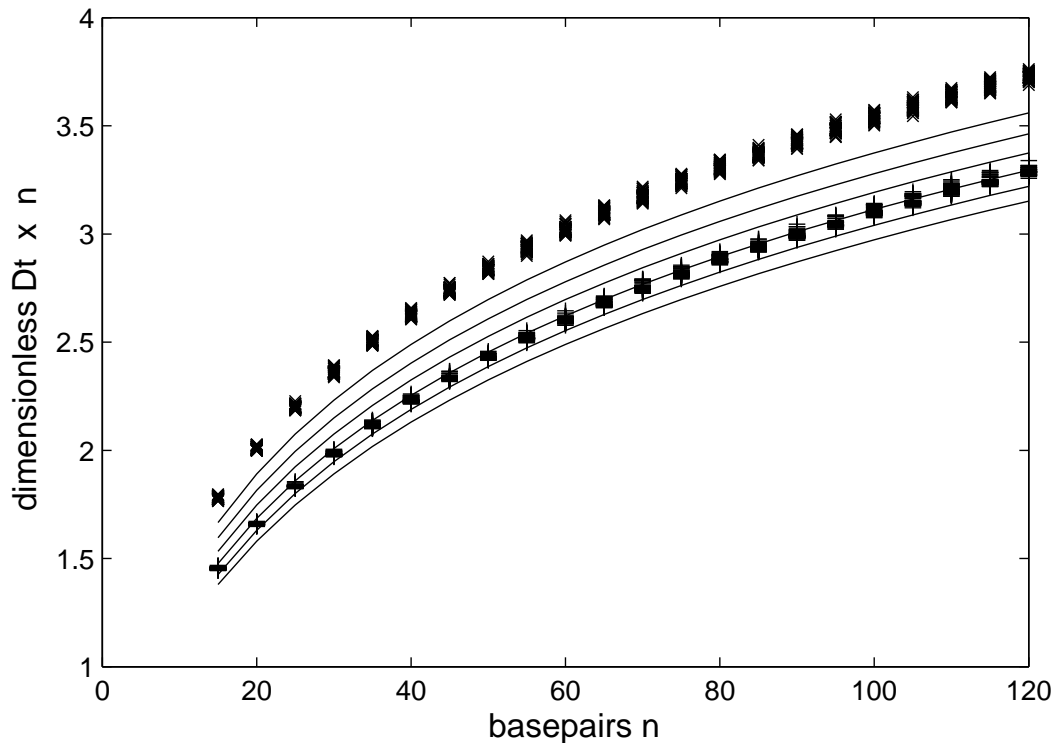

Sequence-dependent model.

Results for idealized model:
Diffusion coefficient vs length

- Symbols: experiments (ultracentrifuge, light scattering, electrophoresis).

- Curves: numerics w/$r = 10, 11, \ldots, 15\text{Å}$ (top to bottom).

- Results predict DNA radius of $r = 10 - 15\text{Å}$.

Results for sequence-dependent model:
Diffusion coefficient vs length

- Curves: straight tube w/$r = 10, 11, \ldots, 15 \mathring{A}$ (top to bottom).

- Crosses: random sequences of prescribed lengths w/$r = 10 \mathring{A}$.

- Pluses: random sequences of prescribed lengths w/$r = 15 \mathring{A}$.

- Results show systematic offset.

- Results now predict DNA radius of $r = 12 - 17 \mathring{A}$.

Closing remarks

- Simplest model of sequence-dependent DNA geometry involves many parameters ($\eta \in \boldsymbol{R}^{26}, \theta \in \boldsymbol{R}^{26}, r \in \boldsymbol{R}$).

- Estimates for all parameters exist, but there is little consensus.

- Hydrodynamic modeling offers interesting approach to validate/refine parameters.

- Successful realization of approach poses many challenges.