Bubble Nucleation And Cooperativity in Duplex Stretching of a Helicoidal DNA Model

Master's Thesis Defense

Pawan Thapaliya

July 2nd, 2019

UTRGV, Department of Physics

Committee Members

Dr. Andreas Hanke

Chair of Committee

Dr. Ahmed Touhami

Committee Member

Dr. Heyongjun Kim

Committee Member

Introduction

1.1 DNA and it's structure

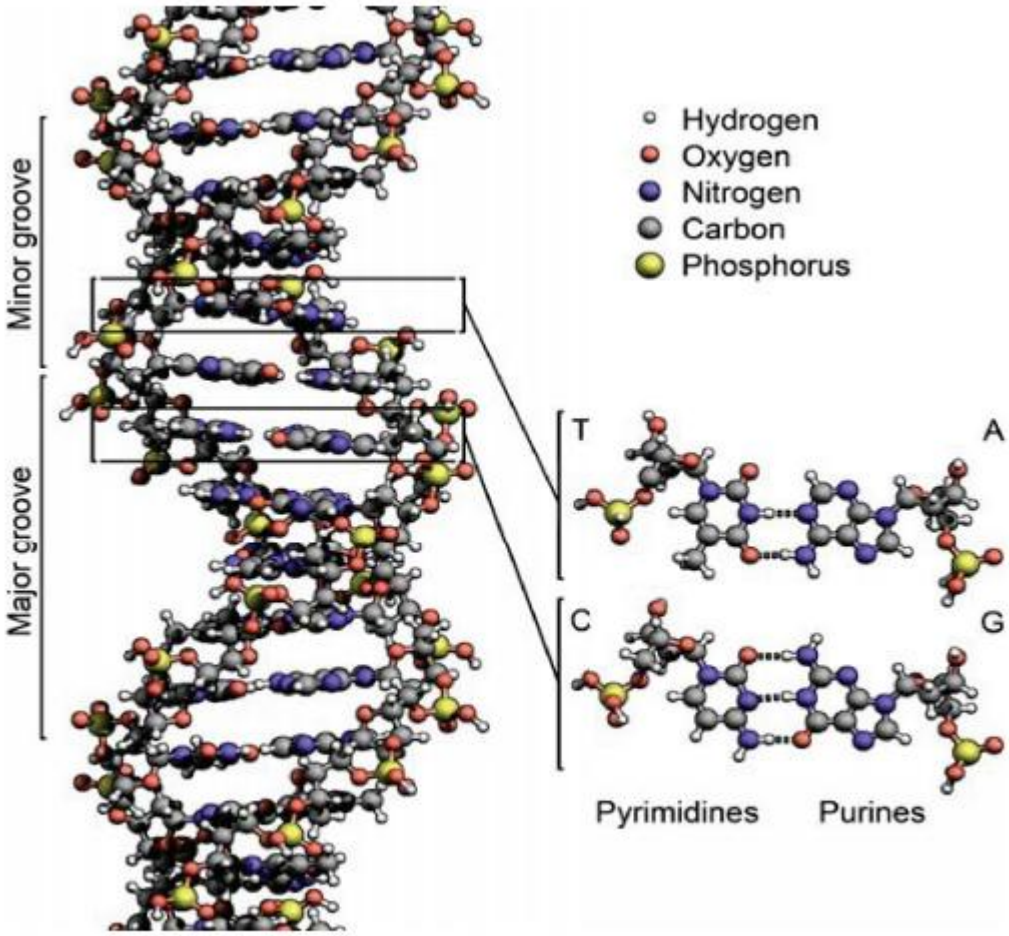


Figure 1.1: Watson Crick Double Helix of DNA helicoid. The figure shows AT base pairs have two hydrogen bonds whereas GC base pairs have three hydrogen bonds.

Basics

- Deoxyribonucleic Acid
- Component sugar, phosphate and Nitrogen bases
- Two strands are complementary to each other start from 5' to 3'
- Bases are paired as AT and GC
- Function carrier of genetic information.

Stability

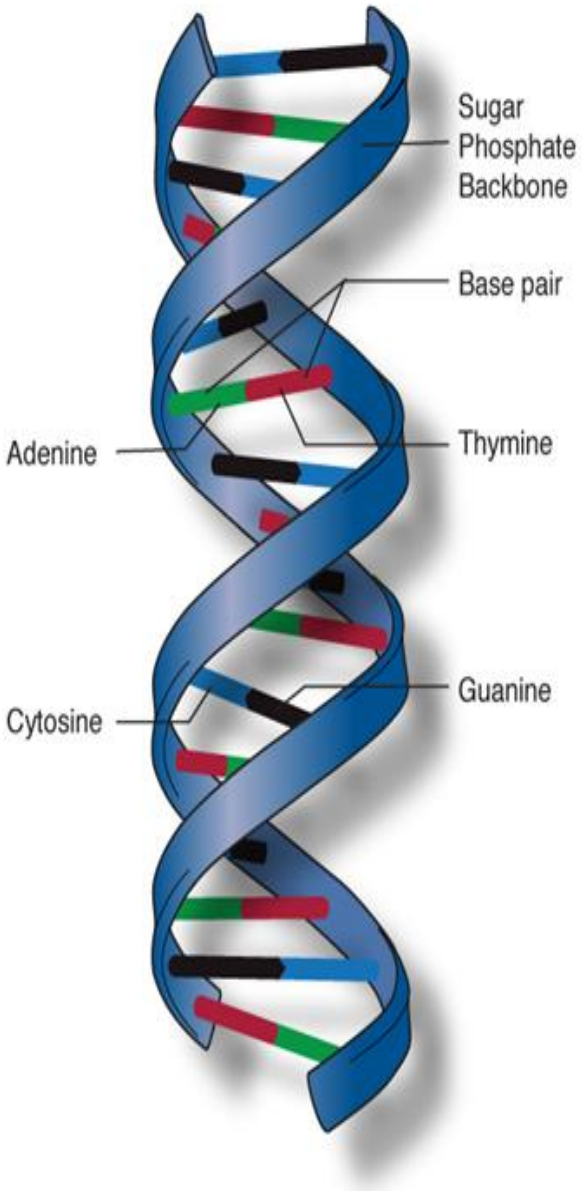
- Covalent bond
- Hydrogen bond
- Stacking interaction hydrophobic effect
- Van der Waals interaction
- dipole-dipole interaction

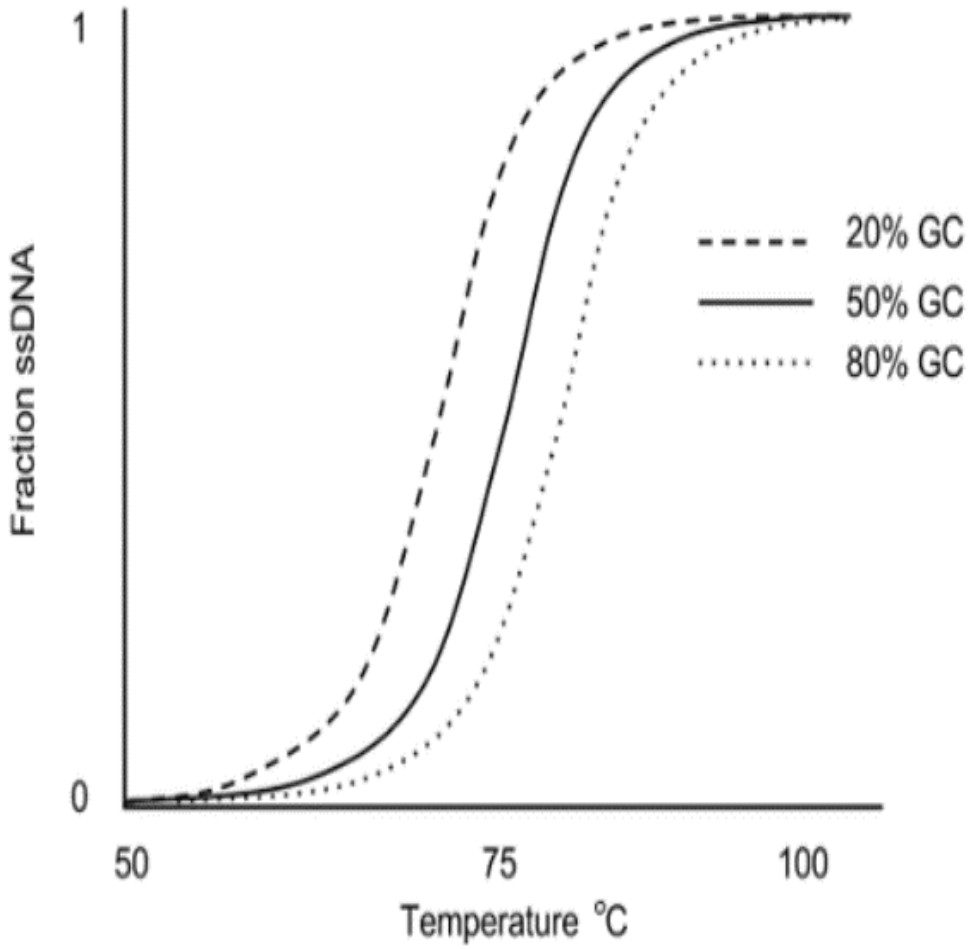
1.2 DNA denaturations by heating and stretching

Thermal DNA melting

M. Peyrard and A. R. Bishop, PRL **62**, 2755 (1989)

T. Dauxois, M. Peyrard, and A. R. Bishop, PRE **47**, 684 (1992)



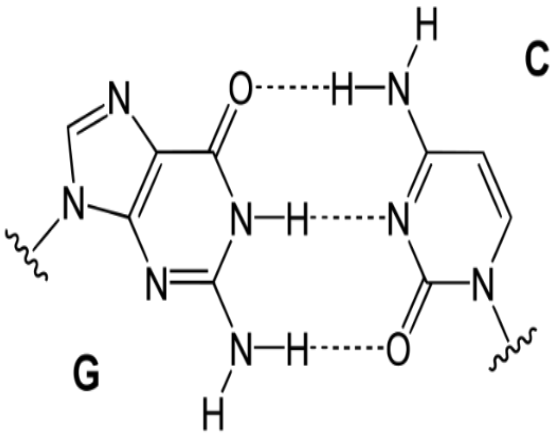


DNA melting curves for three strands of DNA with different levels of GC content. The y-axis indicates the fraction of DNA molecules that are single-stranded.

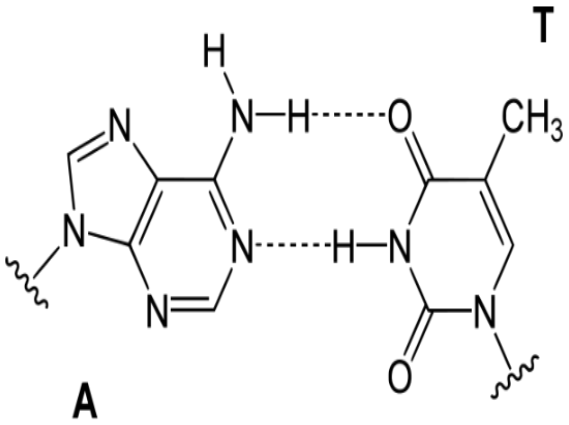
Observation: GC-rich DNA melts at higher temperatures.

Observation: GC-rich DNA melts at higher temperatures.

Reason: GC base pairs are stronger bound than AT base pairs:



GC: 3 H-bonds



AT: 2 H-bonds

Goal: Obtain coarse-grained model to quantitatively predict DNA melting curves for short DNA (10-100 pbs) with given sequence of base pairs

S. Ares, N. K. Voulgarakis, K. O. Rasmussen, and A. R. Bishop, PRL **94**, 035504 (2005)

(a) *Bubble-in-the-middle sequences.*

L60B36: CCGCCAGCGGCGTTATTACATTTAATTCTTAAGTATTATAAGTAATATGGCCGCTGCGCC

L42B18: CCGCCAGCGGCGTTAATACTTAAGTATTATGGCCGCTGCGCC

L33B9: CCGCCAGCGGCCTTTACTAAAGGCCGCT GCGCC

(b) *Bubble-at-the-end sequences.*

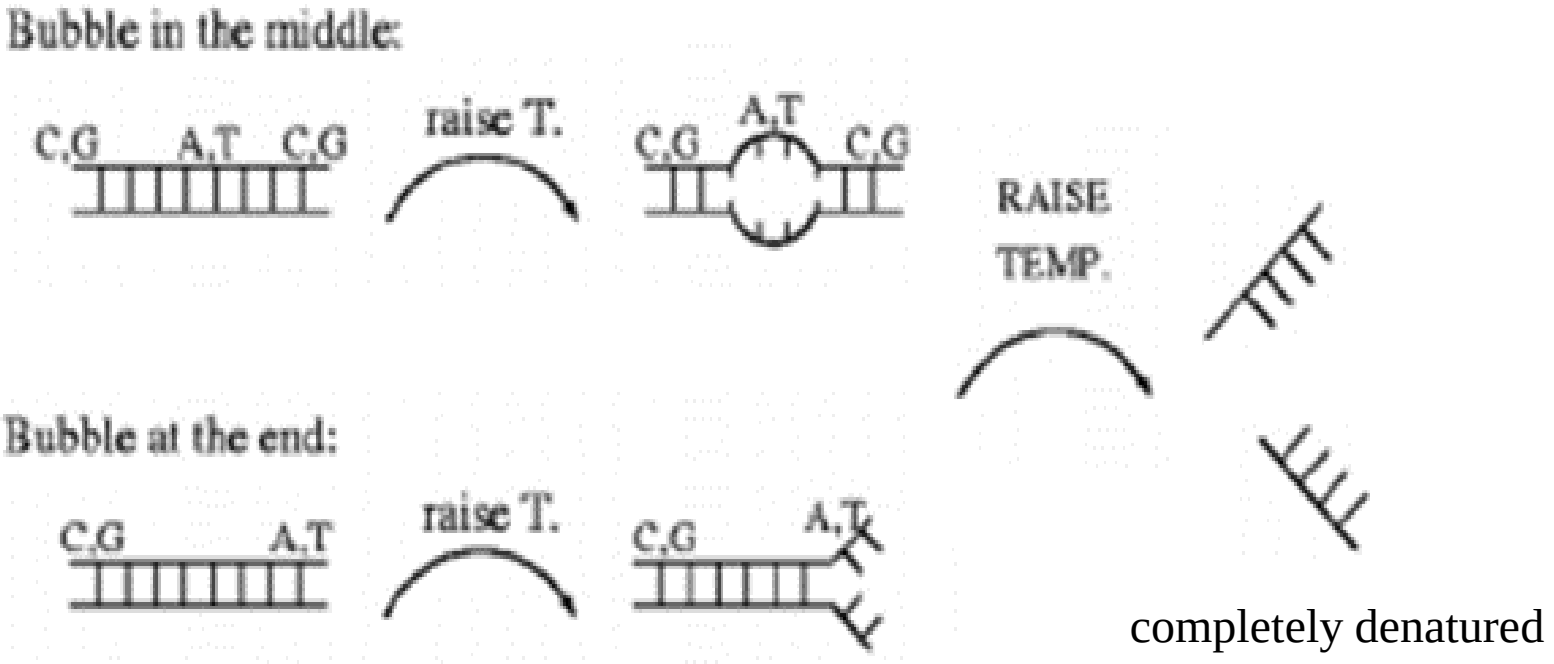
L48AS: CATAATACTTTATATTTAATTGGCGGCGCACGGGACCCGTGCGCCGCC

L36AS: CATAATACTTTATATTGCCGCGCACGCGT GCGCGGC

L30AS: ATAAAATACTTATTGCCGCACGCGTGCGGC

L24AS: ATAATAAAATTGCCCGGTCCGGGC

L19AS2: ATAATAAAGGCGGTCCGCC.



Ares et al. PRL (2005)

B. Denaturations by stretching

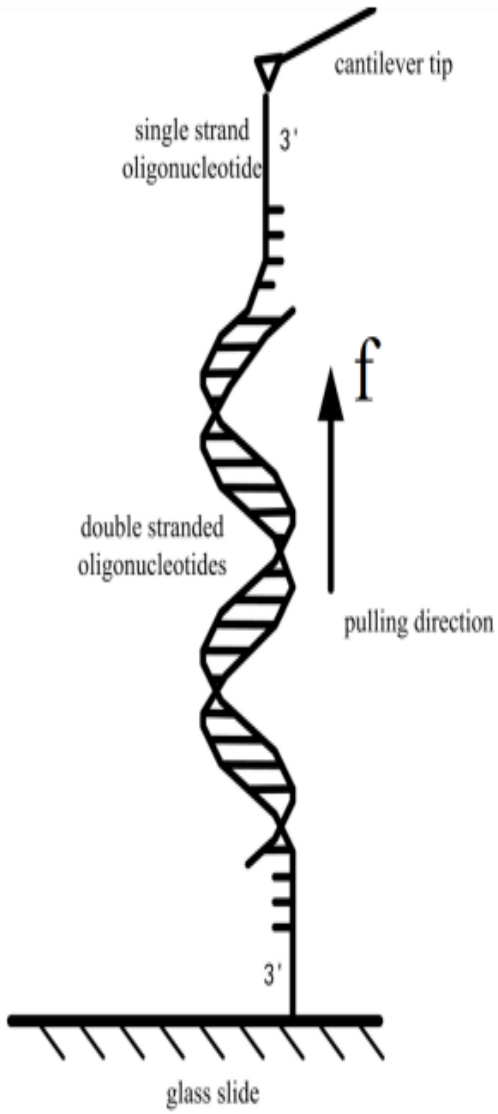


Figure 2.2: Experimental setup for the measurement of force-extension relations of short DNA duplexes using atomic force spectroscopy. Two complementray single strands are covalently immobilized on glass slides and AFM cantilevers. The force f applied to the DNA duplex is measured by the deflection of the AFM cantilever and recorded as a function of the distance between the cantilever tip and the surface.

- Application of mechanical stress to the DNA molecules.
- Optical tweezers and atomic force microscopes enabled the stretching and twisting activities on DNA molecules in the Biophysics experiment

  How?
- The tip of one end of the DNA is chemically anchored to a surface and the sensor attached to the other end measures the applied force.
- In optical or magnetic tweezers instruments the sensor is a microbead
- In atomic force microscope (AFM) the sensor is a microscopic cantilever, as shown in Fig.

Force versus extension curve

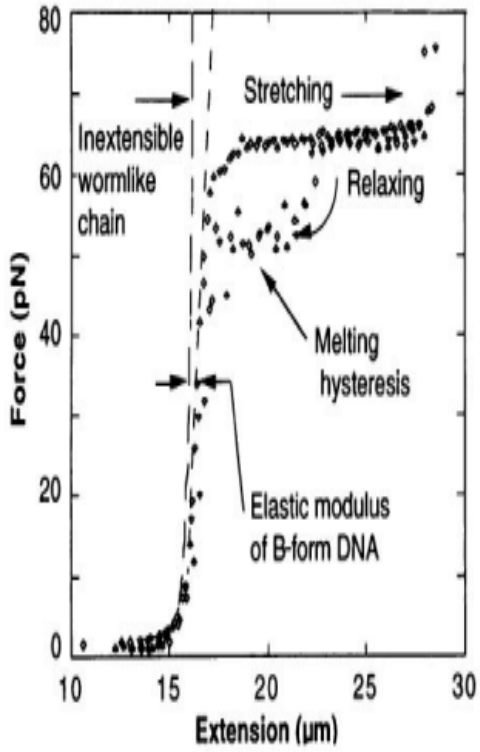


Figure 1.3: Experimental force-extension relation for phage-λ DNA. When the molecule is stretched beyond its B-form contour length it shows a highly cooperative overstertching transition at 65 pN where the DNA extension increases by a factor of 1.7 with very little force increase. The transition is interpreted as force-induced melting where double-stranded DNA is gradually converted to single-stranded DNA as the DNA extension is increased

- understand the mechanical properties of DNA.
- DNA stretching experiments are applicable to study long DNA molecules, such as phage λ DNA (~ 50,000 bp).
- DNA is coupled between t two polystyrene beads and the force F on the DNA is calculated as a function of the DNA extension
- When long DNA molecules are extended beyond their Bform contour length a structural transition occurs during which the extension L of the DNA increases to almost twice its B-form contour length over a very small force range.
- Force extension relations F(L) demonstrates the plateau at forces of Fm ~ 65 pN where the DNA extension per base pair increases from 0.34 to about 0.55 nm .
- At Fm the DNA undergoes a force induced melting transition where double stranded DNA is progressively converted to single stranded DNA,

Method

Peyrard Bishop model for helicoidal DNA

- Study DNA at base pair level.
- opposite bases move in the direction of hydrogen bond
- bases in the same strand are coupled harmonically.

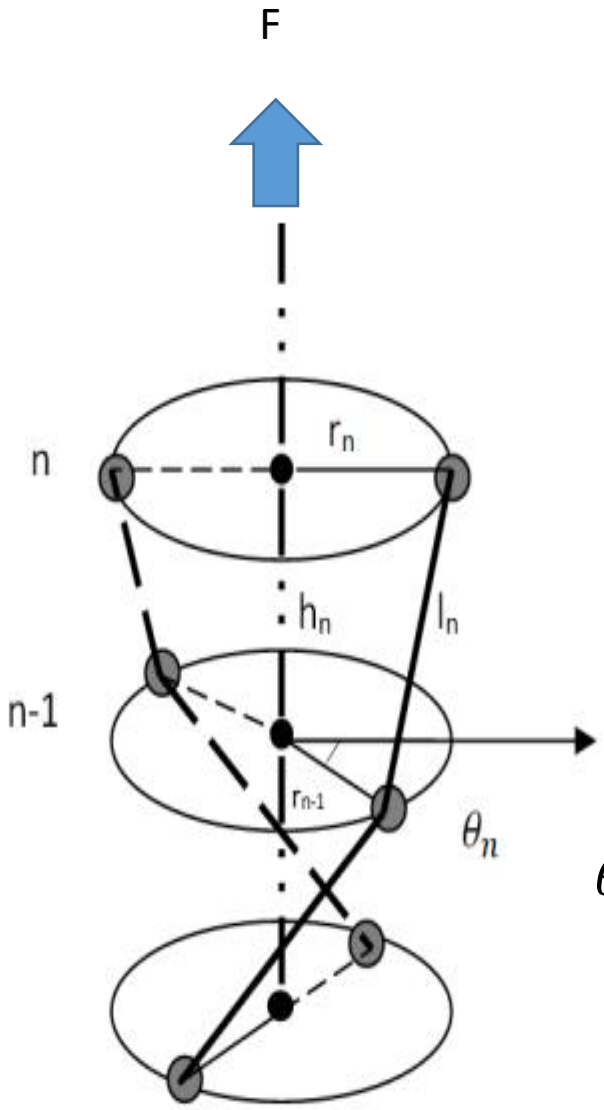
Features

- Local melting of hydrogen bond.
- Formation of denaturations bubbles.

Limitations:

Discuss one dimensional DNA and missed Helicoidal structure of it.

Helicoidal



Require extension of it to Helicodial model

Two types of Helicoidal Model:

1. Fixed base planes and elastic backbones
2. The varying base planes and fixed backbones

address the coupling between opening and twist that results from the helicoidal geometry.

r_n radial displacement

h_n rise along helical axis.

l_n Sugar-phosphate backbone length.

$\theta_n = \phi_n - \phi_{n-1}$ Local twist angle

$\frac{l_n}{h_n} > 1$ for $l_n > h_n$ helicity structure

Figure 2.1: Helicoidal Model: r_n and ϕ_n determines the base pair positions in the plane. Axial dis varies as a function of r_n, θ_n, and l_n

$\{r_n, \theta_n, l_n\} := \{(r_n, \theta_n, l_n), n = 1, \ldots, N\}$. Conformation of DNA molecules

Implement periodic boundary conditions by identifying base pair 0 with base pair N. For example,

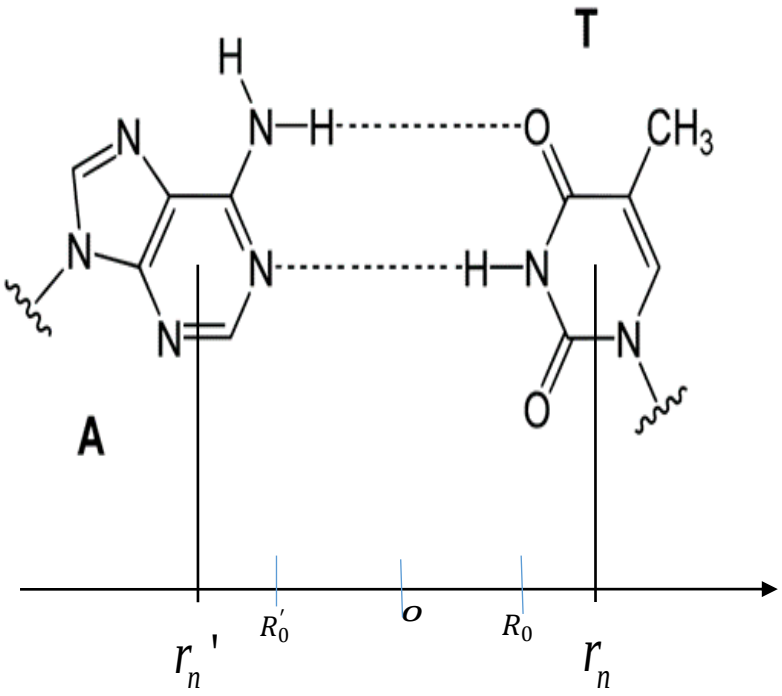
the rise h_1 is calculated using the radial distances r_1 and $r_0 \equiv r_N$ of base pairs 1 and N

$$U\{r_n, \theta_n, l_n\} = \sum_{n=1}^{N} U_m(r_n) + \sum_{n=1}^{N} [U_s(r_n, r_{n-1}) + U_b(l_n) + U_c(r_n, r_{n-1}, \theta_n, l_n) - f h_n]$$

Parameter	Symbol	Value	Unit
Morse potential inverse range	a_n	$a_{\mathrm{AT}} = 8.4$ $a_{\mathrm{GC}} = 13.8$	$\mathrm{\AA}^{-1}$
Morse potential depth	D_n	$D_{\mathrm{AT}} = 0.058$ $D_{\mathrm{GC}} = 1.5\,D_{\mathrm{AT}}$	eV
Equilibrium radial distance of a base	R_0	10	Å
Threshold radial distance to consider base pair open	r_d	10.25	Å
Backbone segment equilibrium length	L_0	6.853	Å
Backbone stretching elastic constant	B	0.128	$\mathrm{ev\,\AA}^{-2}$
Stacking interaction elastic constant	S	0.1	$\mathrm{ev\,\AA}^{-2}$
Stacking interaction coupling parameter	ρ	2	
Stacking interaction inverse range	b_s	0.7	$\mathrm{\AA}^{-1}$
Elastic coupling twist, rise, base pair opening	C	0.028	$\mathrm{ev\,\AA}^{-2}$
Elastic coupling inverse range	b_{el}	0.1	$\mathrm{\AA}^{-1}$
Rise per base pair at mechanical equilibrium	H_0	2.9	Å
Energy constant for twist	c_{tw}	5.64	$\mathrm{eV/rad}^2$

Table 2.1: Set of parameters for the potential energy.

1. Binding energy b/w bases in base pair *n*



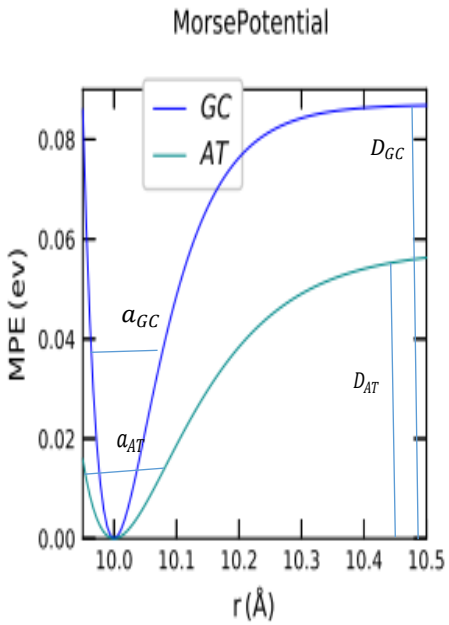


Figure 2.3: Morse potential for AT and GC base pairs. This figure compares range and depth of morse potential for AT and GC base pairs for depth $D_{AT} = 0.058$ and $D_{GC} = 1.5 \times D_{AT}$ and for range $a_{AT} = 8.5\text{Å}^{-1}$ and $a_{GC} = 13.8\text{Å}^{-1}$

R_0 and R_0' is the equilibrium position of the bases

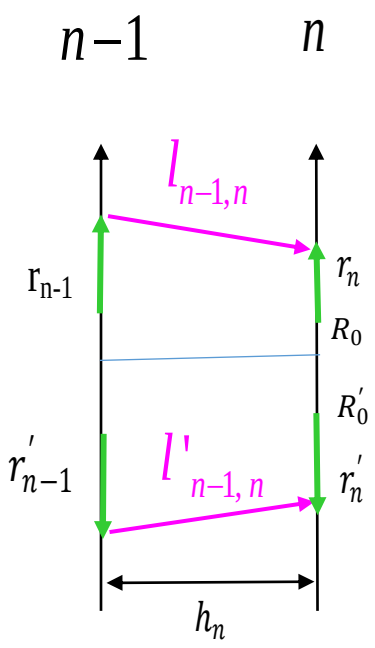
$$U_m(r_n) = D_n\left[e^{-a_n(r_n - R_0)} - 1\right]^2 \quad \text{Morse Potential}$$

D_{AT} = 0.058 eV, D_{GC}= 1.58 D_{AT} eV binding energy of base pairs AT,GC

a_{AT}= 8.4Å^{-1} , a_{GC}=13.8 Å^{-1} range of Morse potential

Ares et al. PRL (2005)

2. Stacking energy b/w base pairs $n-1, n$



$l_{n-1,n}$, $l'_{n-1,n}$ distance b/w bases of adjacent base pairs

h_n distance b/w base pair planes

$x = \frac{1}{\sqrt{2}}(r_n - r_{n-1})$ out-of-phase motion of bases(H-bond stretching)

$y = \frac{1}{\sqrt{2}}(r_n + r_{n-1})$ in-phase(center of mass)

→ Stacking energy:

$$U_s(r_n, r_{n-1}) = \frac{s}{2}(1 + \rho e^{-b_s(r_n + r_{n-1} - 2R_0)})\,(r_n - r_{n-1})^2$$

$$k(r) = 1 + e^{-b_s \Delta(r)} \rightarrow \begin{cases} 1 + \rho = 3, \Delta(r) = 0 \;\; bps\ closed \\ 1, \Delta(r) \rightarrow \infty \text{ bps open} \end{cases}$$

S= 0.1 eVÅ^{-2}

b_s = 0.7 Å^{-1}

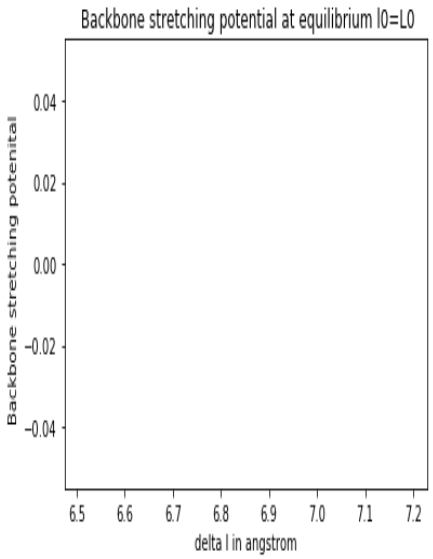
$\rho = 2$

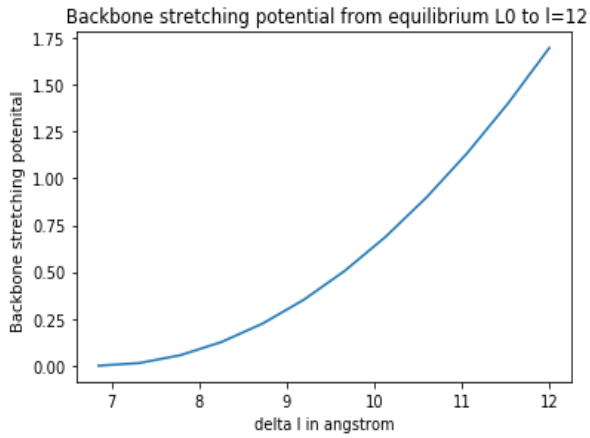
Switch function: stacking interaction b/w closed bps is stronger

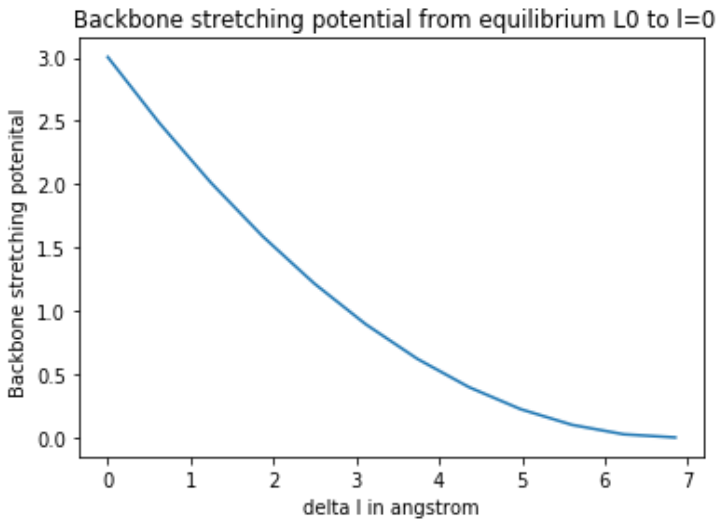
Ares et al. PRL (2005)

Backbone stretching energy b/w n-1,n bases

$$U_b(l_n) = \frac{1}{2} B (l_n - L_0)^2$$







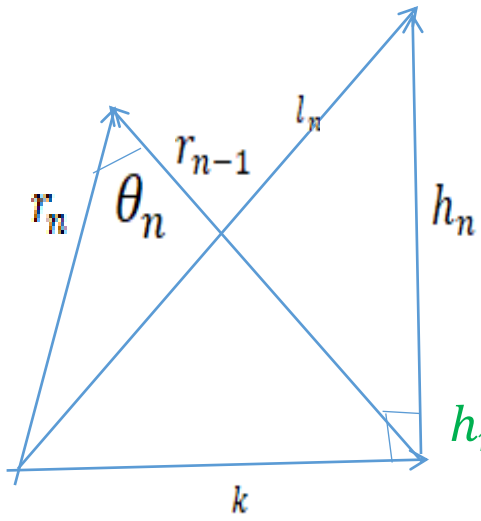
Coupling Between helical twist, rise, and base pair opening

$$U_c(r_n, r_{n-1}, \theta_n, l_n) = \frac{c_{tw}}{2} e^{-bc\,(r_n + r_{n-1} - 2R_0)}(h_n - H_0)^2$$

Describe cross over from closed to open

$$k + r_{n-1} = r_n$$

$$k = r_n - r_{n-1}$$



$$k^2 = r_n^2 + r_{n-1}^2 - 2r_n r_{n-1}\cos\theta_n$$

$$h_n^2 + k^2 = l_n^2$$

$$h_n = \sqrt{l_n^2 - r_n^2 - r_{n-1}^2 + 2r_n r_{n-1}\cos\theta_n}$$

$$h_n = \sqrt{l_n^2 - 2r_n^2(1 - \cos(\theta_n)},\ r_n = r_{n-1}$$

$$h_n = \sqrt{l_n^2 - (r_n - r_{n-1})^2},\ \theta_n = 0.$$

h_n is bounded by $h_n \in [0, l_n]$

θ_n is bounded by $\theta_n \in [-0.1, 0.7]$

Contour plot h and U_c

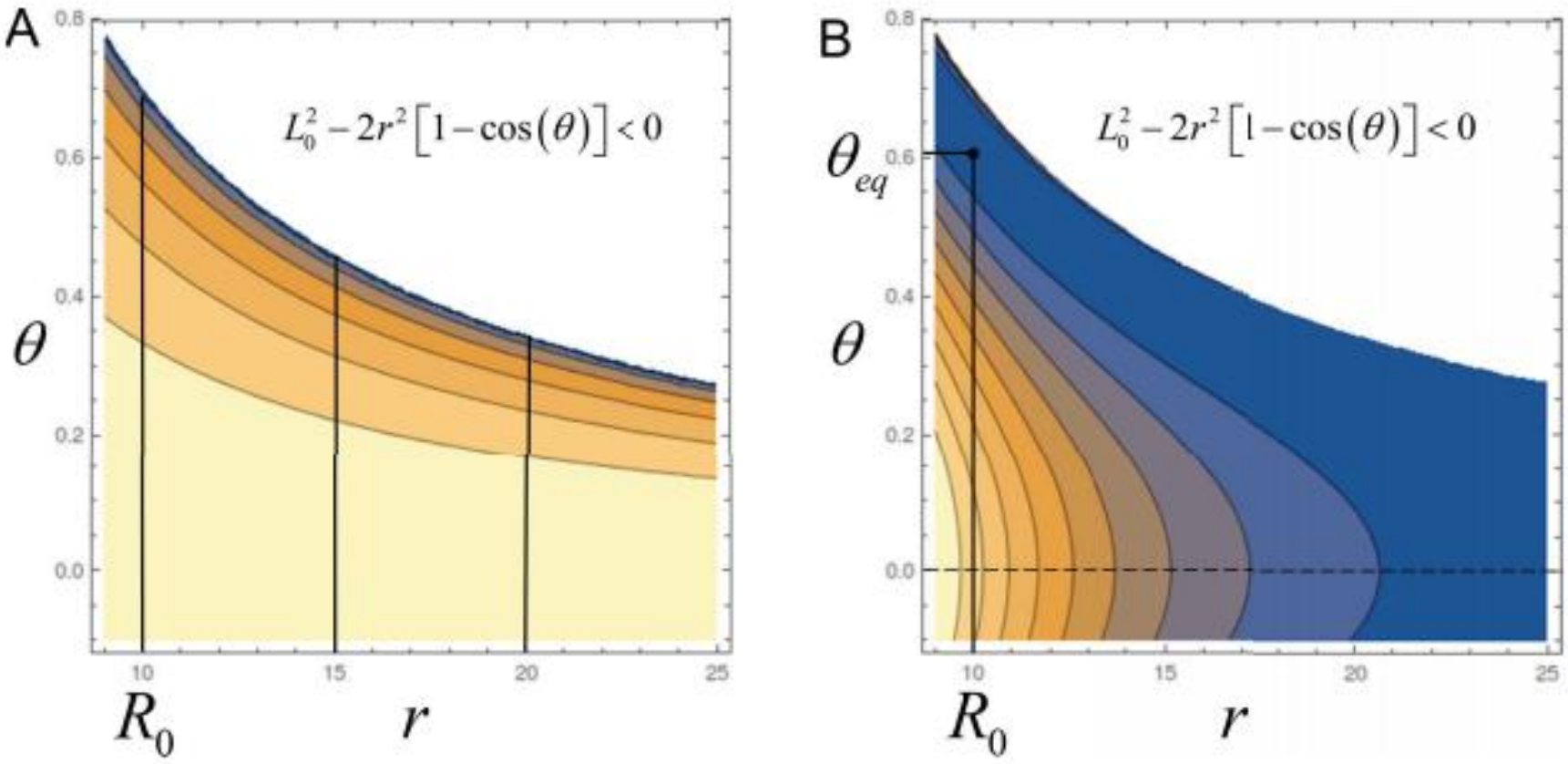


Figure 2.4: Contour plots of (A) the helical rise $h_n(r,\theta)$ and (B) corresponding elastic energy $U_c(r,\theta)$ as function of $r_n = r_{n-1} \equiv r$ and $\theta_n \equiv \theta$ for fixed $l_n = L_0$ (see Eqs. (2.6) and (2.8)). The upper right region (shown in white) correspond to negative values of the argument of the square root in Eq. (2.8) and is excluded. (A) shows that for fixed radial distance r (vertical lines) an increasing helical rise h corresponds to decreasing twist θ, *i.e.*, unwinding of the double helix. The shape of the minimum of the elastic energy $U_c(r,\theta)$ in (B) shows that a decreasing twist θ, in turn, favors increasing values r, *i.e.*, opening of base pairs.

External stretching force between n and n-1.

$$U_f(r_n, r_{n-1}, \theta_n, l_n) = -f h_n$$

- f acts on the whole terminal base pair simultaneously, corresponding to a force f /2 applied to each base of the terminal base pair.
- Terminal base pair is free to rotate about the helical axis.
- f favors unwinding of the double helix and opening of base pairs.

2.2 Monte Carlo Simulation

$$U\{r_n, \theta_n, l_n\} = \sum_{n=1}^{N} U_m(r_n) + \sum_{n=1}^{N} [U_s(r_n, r_{n-1}) + U_b(l_n) + U_c(r_n, r_{n-1}, \theta_n, l_n) - f h_n]$$

Goal: Obtain melting curves as a function of temperature for thermal denaturations, and as a function of temperature and force for force-induced denaturations .

- average fraction f of open base pairs from the condition$\langle r_n \rangle > r_{open}$
- average fraction l of open base pairs from the conditions at least one base pairs is not open
- average number of the bubbles n_b formed for overall simulation
- average size of the bubble per base pairs l_b formed for overall simulation.
- average rise of base plane distance h for over all simulation
- average twist angle θ for overall simulation

Method: Obtain the profile $\langle r_n \rangle, \langle \theta_n \rangle$ for given sequence of bps $n = 1, \ldots N$ by MC simulation; calculate l by applying the condition that at least one base pairs is closed and f by using the criterion that bp n is considered "open" if $\langle r_n \rangle \geq 10.25$Å.

Algorithm:We use the standard metropolis Monte Carlo algorithm to produce an equilibrium state of the system. Initialize base pairs , pick single base pair n_0 at random and draw new value of the variable $r_{n\,0}, l_{n\,0}$, $and\ \theta_{n\,0}$ from random Gaussian distribution. The energy of the old and new configuration is calculated. If $U_{trl} < U_{current}$,the new configuration is accepted unconditionally. Otherwise new configurations accepted with Metropolis probability for $U_{trl} > U_{current}$

$$p = min\left(1, \exp^{-\frac{U_{trl} - U_{cur}}{k_{BT}}}\right)$$

Ares et al. PRL (2005)

Base-pair opening move

$U_s(r_n, r_{n-1}) = \frac{S}{2}(r_n - r_{n-1})^2$ stacking interaction b/w r_n&r_{n-1} bases.

$x = \frac{1}{\sqrt{2}}(r_n - r_{n-1}), y = \frac{1}{\sqrt{2}}(r_n + r_n - 1)$ the out of phase and in phase(center of mass) relative motion of $r_n and\ r_{n-1}$

in matrix form

$$\begin{pmatrix} x \\ y \end{pmatrix} = A \begin{pmatrix} r_n \\ r_{n-1} \end{pmatrix}$$

det(A) = 1,

i.e., the determinant of the Jacobian matrix A of the transformation (r_n r_{n-1}) → (x, y) is equal to one.

In terms of the variable x ,

$Us(x) = S\ x^2$

$\sigma = \sqrt{\langle x^2 \rangle} = \sqrt{\frac{k_B T}{2S}}$ is the standard deviation of x in a canonical ensemble with Boltzmann weight exp[−βUs(x)]

where $\beta = 1/(k_B T)$ and k_B is the Boltzmann constant.

trial value :

$$r_{trl} = r_{cur} + \sqrt{\frac{k_B T}{2S}}\,\xi$$

where ξ is a standard Gaussian random variable with $\langle \xi \rangle = 0$ and $\langle \xi^2 \rangle = 1$.

$\langle (r_{trl} - r_{cur})^2 \rangle = k_B T / 2S \langle \xi^2 \rangle = k_B T / 2S = \sigma^2$.

variance of base pair opening moves corresponds to the variance of thermal fluctuations of x,

Twisting move

$U_{tw} = \frac{2\pi^2 K}{H_0}(\Delta T_w)^2$ potential energy for twisting

$\Delta Tw = \frac{1}{2\pi}(\theta_n - \theta_0)$ angular displacement per unit twist.

$c_{tw} = \frac{K}{H_0} = 5.64\ ev\ rad^2$ twisting constant obtained from the twist modulus equivalent to $2.61 \times 10^{-19}\ erg\ cm\ per\ rad^2$. and $\theta_n = (\phi_n - \phi_{n-1})$ is angle of twist.

$\sigma = \sqrt{\frac{1}{c_{tw}\beta}}$ Standard deviation σ of θ in a canonical ensemble

$\beta = 1/(k_B T)$ Boltzmann factor

k_B Boltzmann constant.

$\theta_{trl} = \theta_{cur} + \sqrt{\frac{k_B T}{c_{tw}}}\ \xi$ Trial value for twisting move

ξ is a standard Gaussian random variable with $\langle \xi \rangle = 0$ and $\langle \xi^2 \rangle = 1$.

$$\langle (\theta_{trl} - \theta_{cur})^2 \rangle = \frac{k_B T}{c_{tw}} \langle \xi^2 \rangle = \frac{k_B T}{c_{tw}} = \sigma^2 .$$

Stretching Move

$U_b = \frac{1}{2} B(l_n - l_0)^2$ potential for backbone stretching.

$B = K_{mod}/H_0 = 0.128eV/\text{Å}^2$ stretching constant obtained from the stretching modulus equivalent to 594pN per helical rise H_0.

$\sigma = \sqrt{\frac{k_B T}{B}}$ standard deviation of thermal fluctuation

$\beta = 1/(k_B T)$ Boltzmann factor

k_B the Boltzmann constant.

$l_{trl} = l_{cur} + \sqrt{\frac{k_B T}{B}}\xi$ Trial values for stretching move

ξ is a standard Gaussian random variable with $\langle \xi \rangle = 0$ and $\langle \xi^2 \rangle = 1$.

$\langle (l_{trl} - l_{cur})^2 \rangle = \frac{k_B T}{B} \langle \xi^2 \rangle = \frac{k_B T}{B} = \sigma^2$.

Double stranded ensemble

dsDNAE is defined as all configuration assigned as dsDNA together with their corresponding Boltzmann weight. If $R\ (\ r^N\ )$ be any function,

$\langle R\rangle \equiv \frac{\int dr^N R(r^N) exp\,[-\beta U\,]}{\int dr^N exp\,[-\beta U\,]}$ statistical average in the full phase space

with $dr^N \equiv dr^N, dr_{n-1}, \ldots dr1.$

The exponential term represents the probability distribution density with Boltzmann factor β.

$\mu(r^N)$ weight function

$\langle R(r^N)\rangle_\mu \equiv \frac{\left\langle R(r^N)_\mu \right\rangle}{\langle \mu \rangle}$ the ensemble average of $R(r^N)$ in dsDNAE

$\mu \equiv 1 - \Pi_{k=1}^{N} \eta_n$

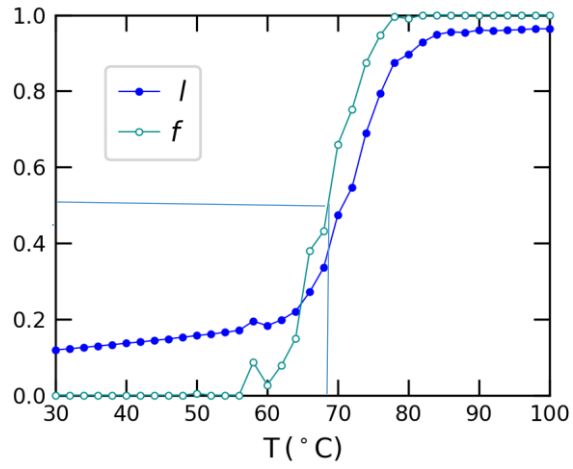
Advantage:

- Removes the problems of nonnormalizability of the full phase space equilibrium distribution
- Represents the actual experimental situations at a temperature below the denaturation transition.

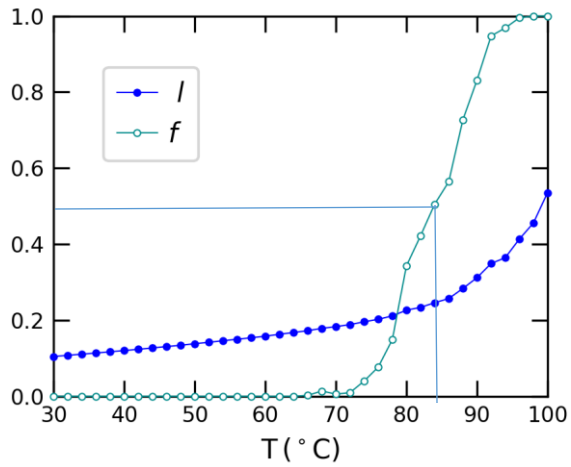
Result

Thermal denaturations

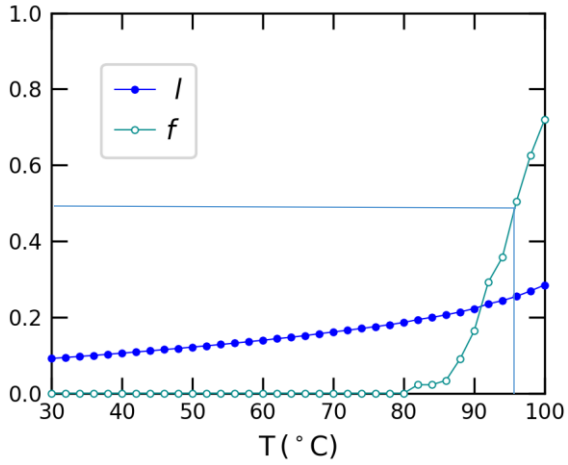
L60B36

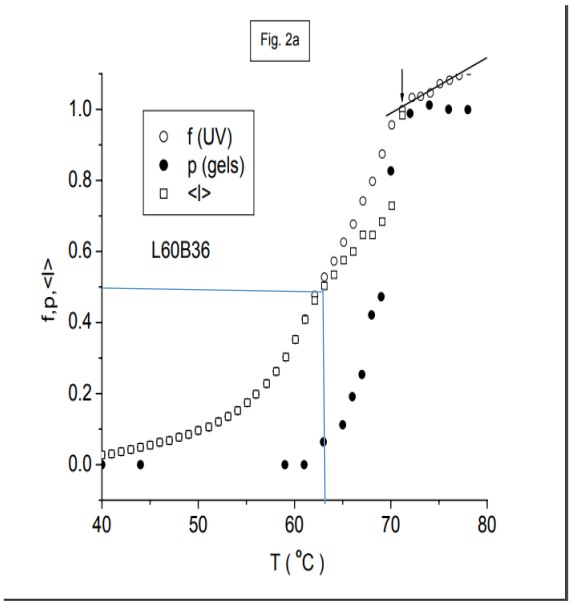


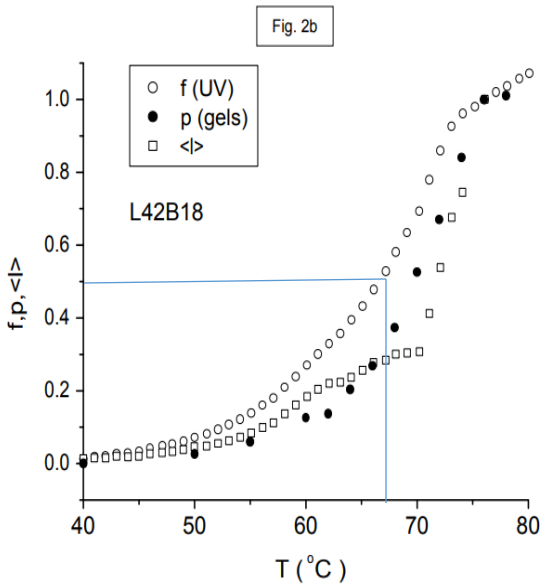
L42B18

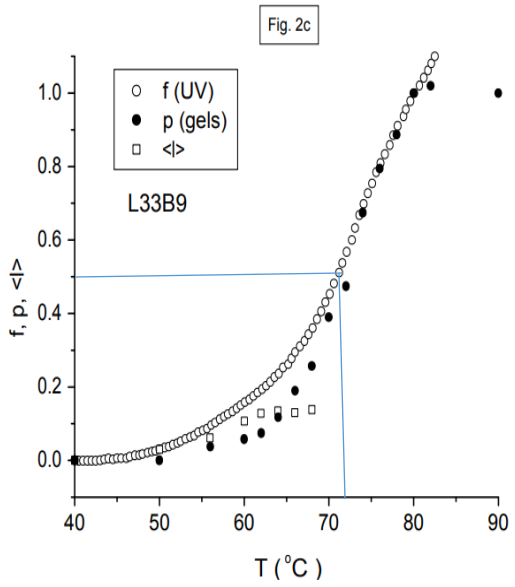


L33B9

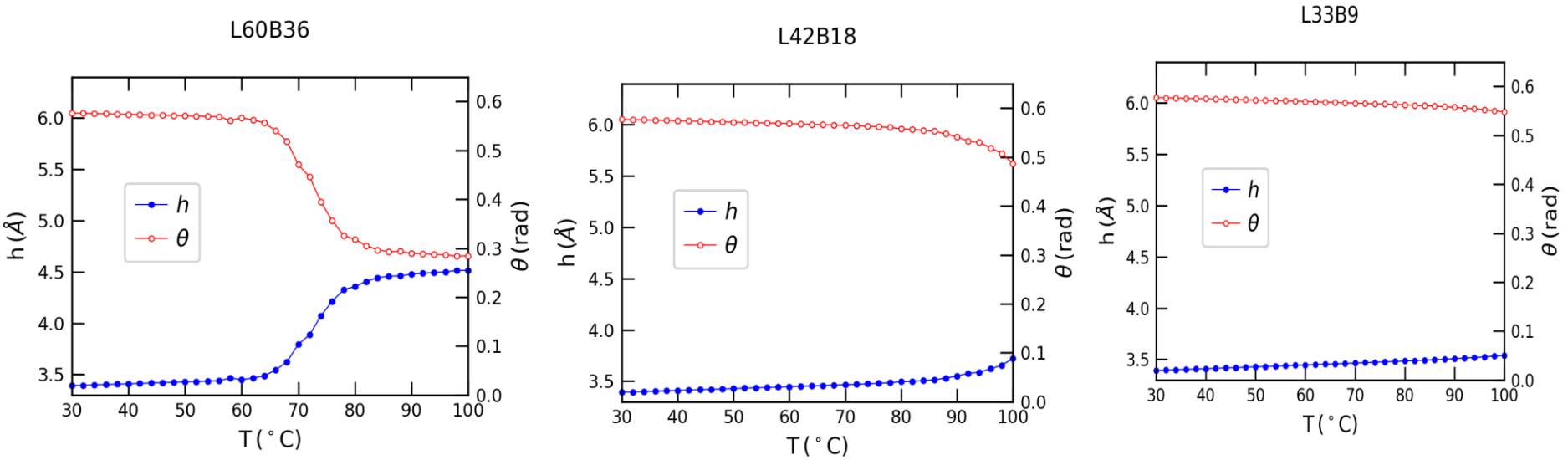




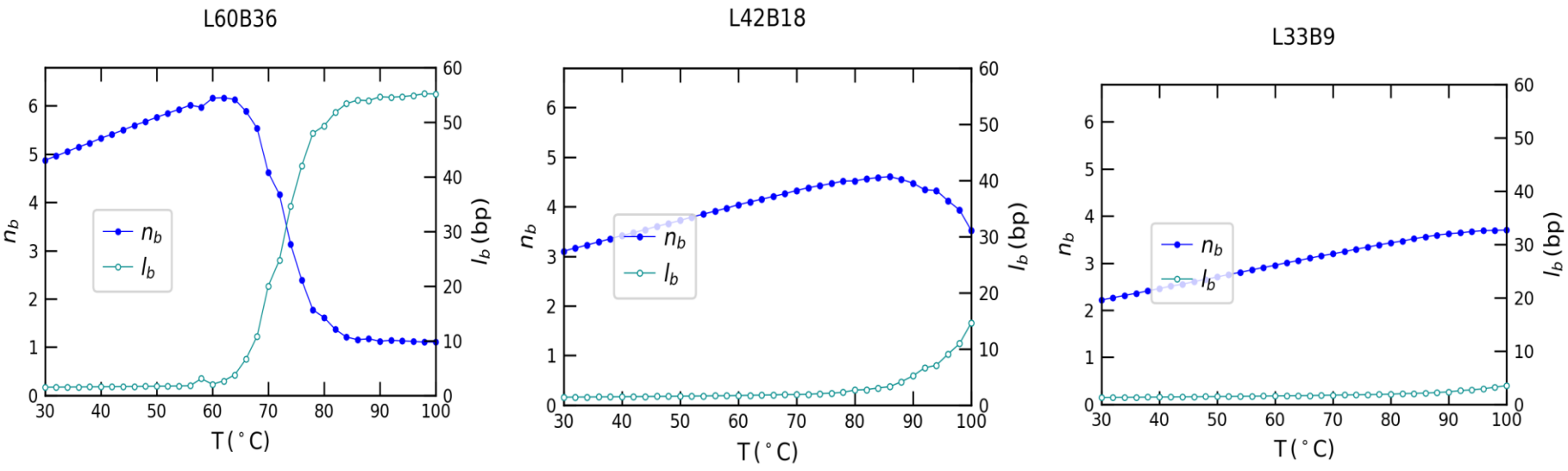




- f represents the average fraction of open base pairs obtained from the condition $\langle r_n \rangle > r_{open}$
- l represents the average fraction of open base pairs obtained from the overall simulation



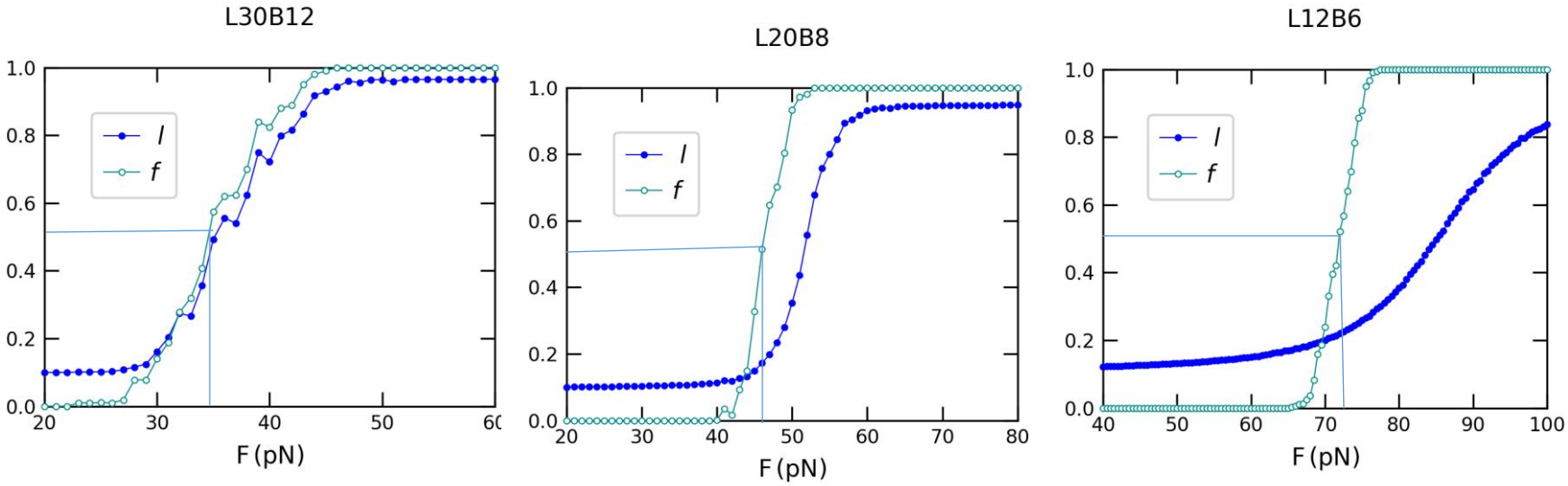
- h represents the average helical rise per base pair for over all simulation
- θ represents the average twist angle per base pair for over all simulation



n_b mean number of bubbles averaged over all simulations

l_b mean bubble sizes averaged over all simulations

Force-induced denaturations



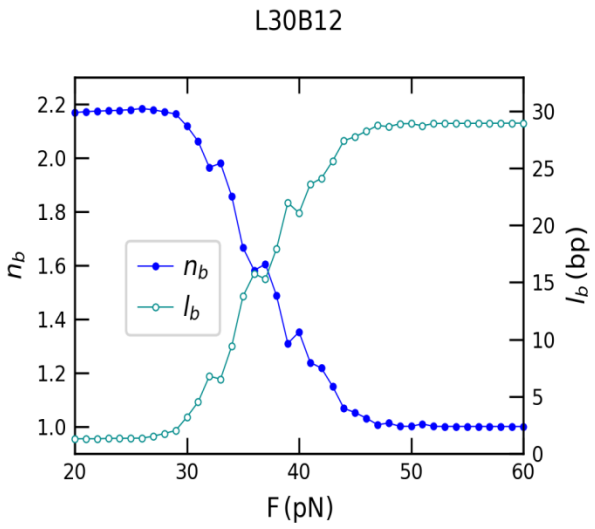
- f represents the average fraction of open base pairs obtained from the condition $\langle r_n \rangle > r_{open}$
- l represents the average fraction of open base pairs obtained from the overall simulation

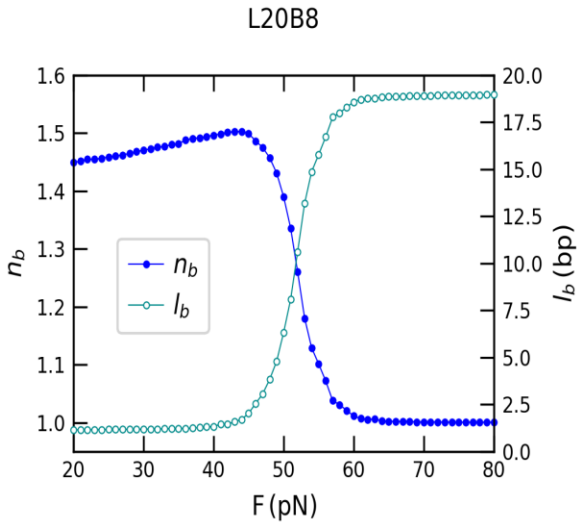
Bubble-in-the middle sequence

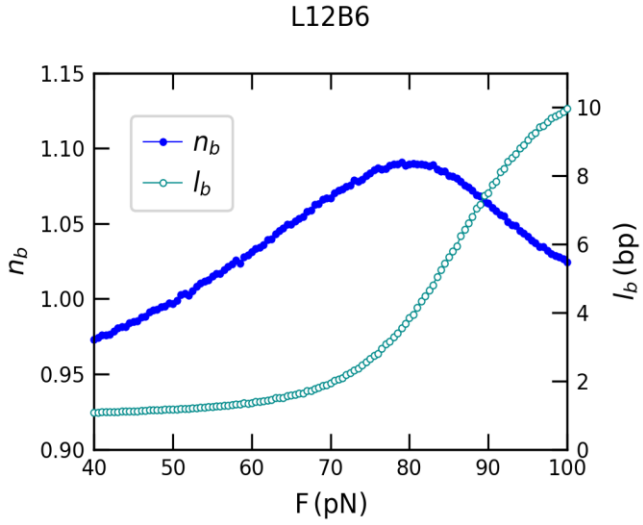
L30B12 GGCTCCCTTCTACCACTGACATCGCAACGG

L20B9 CGTTGGTGCGGATATCTCGG

L12B6 CGCAAAAAAGCG

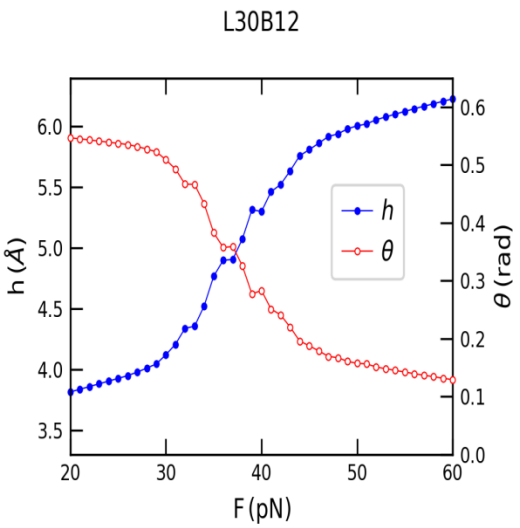


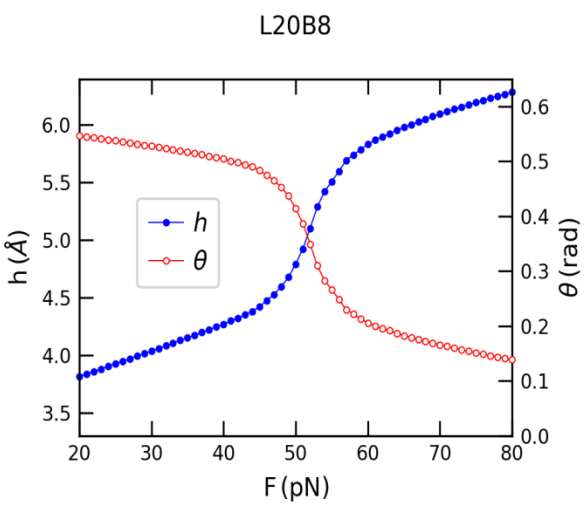


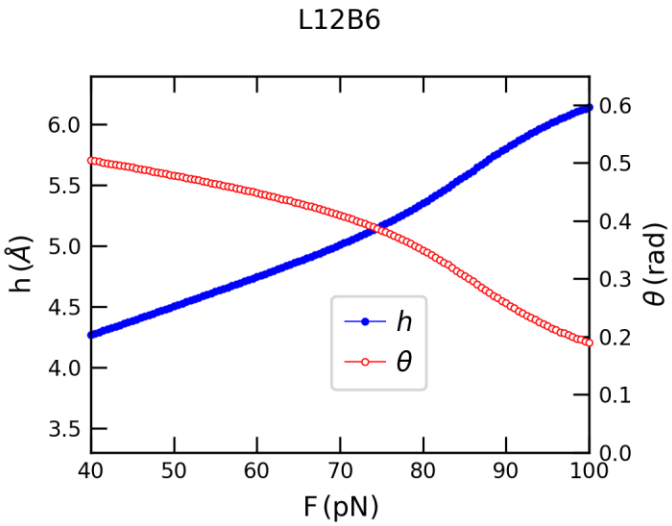


n_b mean number of bubble formed average over all simulation

l_b mean number of bubble size average over all simulation



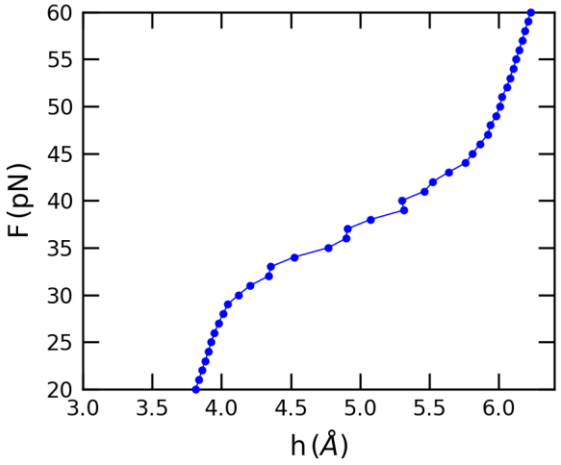




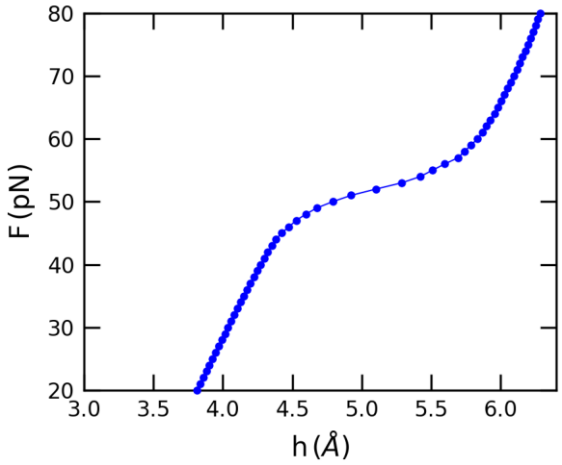
- h represents the average helical rise per base pair for over all simulation
- θ represents the average twist angle per base pair for over all simulation

Force vs. extension

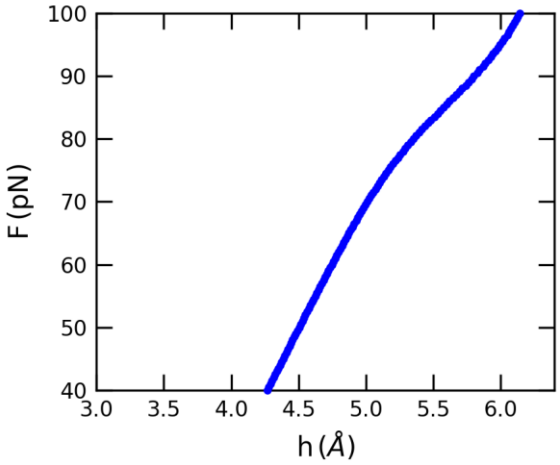
L30B12

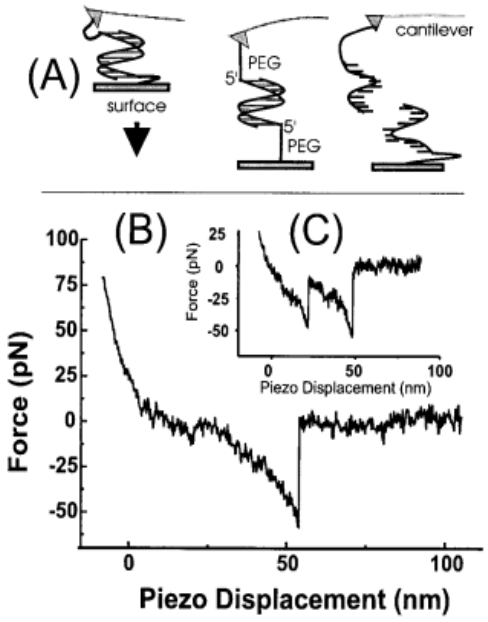


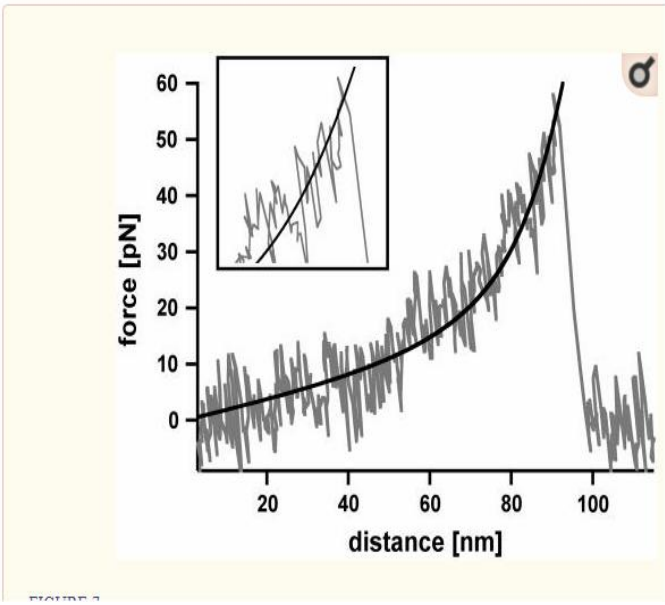
L20B8

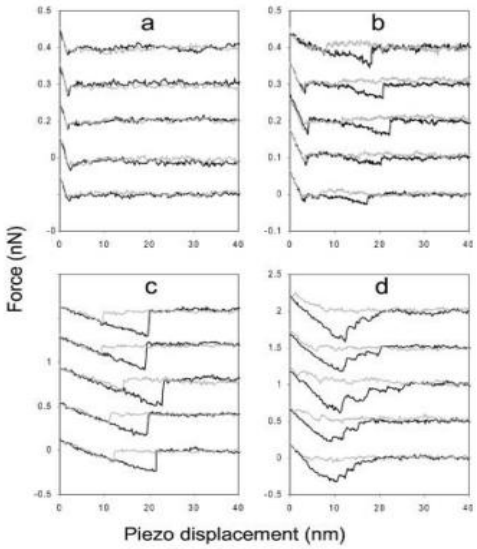


L12B6









Conclusion

- B-S transition cannot be observed in finite short DNA sequence.
- Stress bearing capacity of the DNA increases with the decrease in the sequence of the base pairs.
- Bubble formation and cooperativity of the DNA decreases with the shortening of the DNA sequence.
- Local unwinding of the DNA helix and opening of the base pairs are the characteristics of the force induced DNA denaturations.

Thank you