REPORT DOCUMENTATION PAGE

AFRL-SR-AR-TR-07-0011

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	Dec. 5, 2006	August 15, 2005-August 14, 2006

4. TITLE AND SUBTITLE
DNA conformational dynamics and protein binding probed by DNA overstretching

5. FUNDING NUMBERS
FA9550-05-01-0472

6. AUTHOR(S)
Andreas Hanke

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Texas at Brownsville /
Texas Southmost College

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR 875 N. Randolph St., Ste. 325,
RM 3112
Arlington, VA 22203-1768
Dr. Harold Weinstock/NE

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; distribution is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*
Recent advances in single-molecule force spectroscopy of DNA make it possible to study the thermodynamics and kinetics of DNA binding proteins under a wide range of conditions. A biophysical model for the DNA binding T4 gene 32 protein has been developed to study the kinetics of DNA protein binding to transient single-stranded DNA regions due to thermal fluctuations. The model is used to analyze recent single-molecule spectroscopy data of this system.

14. SUBJECT TERMS	15. NUMBER OF PAGES
	8
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
		Unclassified	

Technical report: DNA Conformational Dynamics and Protein Binding

A. Physiological background of DNA denaturation

The following Introduction is taken from the previous technical report "Regulation of DNA metabolism by DNA-binding proteins probed by single molecule spectroscopy". We repeat this Introduction here for completeness.

The Watson-Crick double-helix is the thermodynamically stable configuration of a DNA molecule under physiological conditions (normal salt and room/body temperature). This stability is effected (a) by Watson-Crick H-bonding, that is essential for the specificity of base pairing, i.e., for the key-lock principle according to which the nucleotide Adenine exclusively binds to Thymine, and Guanine only to Cytosine. Base-pairing therefore guarantees the high level of fidelity during replication and transcription. (b) The second contribution to DNA helix stability comes from base-stacking between neighboring base pairs: through hydrophobic interactions between the planar aromatic bases, that overlap geometrically and electronically, the base-pair stacking stabilizes the helical structure against the repulsive electrostatic force between the negatively charged phosphate groups located at the outside of the DNA double-strand. While hydrogen bonds contribute only little to the helix stability, the major support comes from base-stacking.

The capability of the two complementary strands of DNA to move apart and to rejoin, without damaging the chemical structure of the two single-strands, is crucial to many physiological processes such as replication via the proteins DNA helicase and polymerase, and transcription through RNA polymerase. During these processes, the proteins unzip a certain region of the double-strand, to obtain access to the genetic information stored in the bases in the core of the double-helix. This unzipping corresponds to breaking the hydrogen bonds between the base-pairs. Classically, the so-called melting and reannealing behavior of DNA has been studied in solution in vitro by increasing the temperature, or by titration with acid or alkali (**Figure 1**). During thermal melting, the stability of the DNA duplex is related to the content of triple hydrogen-bonded G-C base-pairs: the larger the fraction of G-C pairs, the higher the required melting temperature or pH value. Conversely, molten, complementary chains of single-stranded DNA (ssDNA) begin to reassociate and eventually reform the original double-helix under incubation at roughly $25°$ below the melting temperature T_m.

Complementary to thermal or pH induced denaturation, ssDNA can be driven toward denaturation mechanically, by applying a tensional stress along the DNA in an optical tweezers trap. The force per extension increases in worm-like chain fashion, until a plateau at approximately 65 pN is reached (**Figure 2**). This plateau is sometimes interpreted as new DNA configuration, the S form. By a series of experiments, it appears more likely that the plateau corresponds to the mechanical denaturation transition. To first order, the effect of the longitudinal pulling with a force F translates into an external torque $\tau = \tau(F)$, whose effect is to decrease the free energy of melting ΔG:

$$\Delta G(F) = \Delta G(F=0) - \tau(F)\,\theta_0 \qquad (1)$$

where $\theta_0 = 2\pi/10.35$ is the twist angle per bp of the double helix without applied forces.

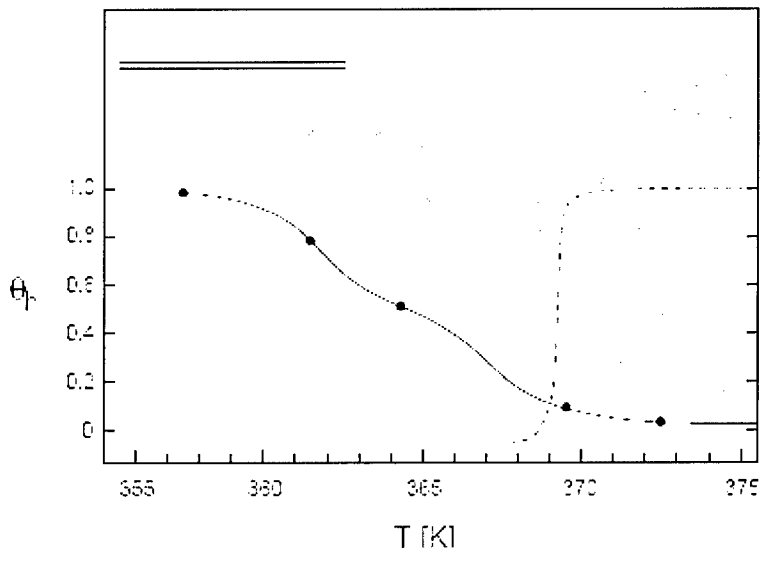

Figure 1. Fraction θ_h of double-helical domains within the DNA as a function of temperature T. Schematic representation of $\theta_h(T)$, showing the increased formation of bubbles and unzipping from the ends until full denaturation has been reached.

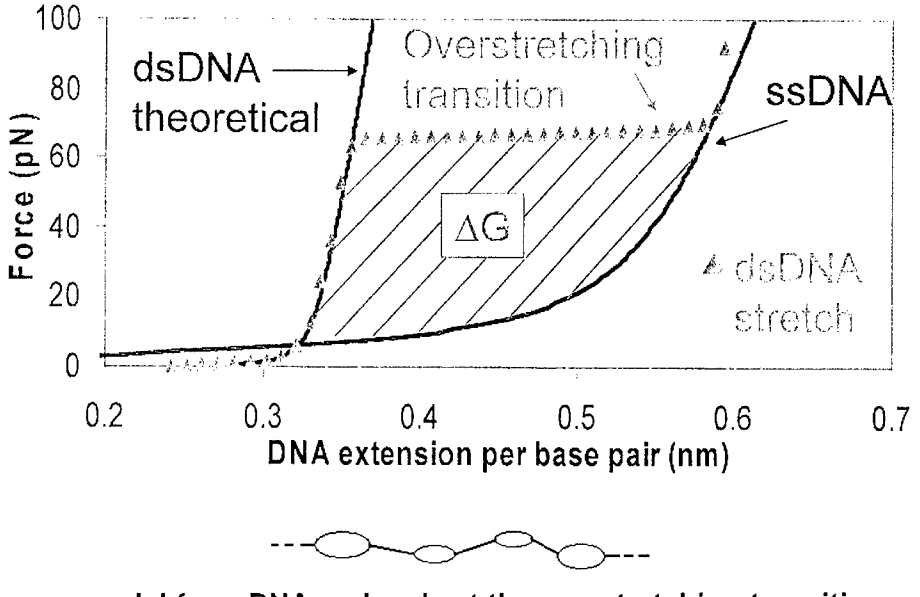

Figure 2. Top: Overstretching transition of double-stranded DNA. The force extension relation exhibits a rapid worm-like chain increase until at around 65 pN a plateau is reached. The transition on the plateau corresponds to progressive mechanical denaturation from dsDNA (left end) to ssDNA (right end). Bottom: Schematic depiction of DNA conformation on the overstretching plateau with local denaturation zones (bubbles). See text for details. Figure courtesy Mark C. Williams.

While the double helix is the thermodynamically stable configuration of the DNA molecule below the melting temperature T_m (at nondenaturing pH), even at physiological conditions there exist local denaturation zones, so-called DNA-bubbles, predominantly in A-T-rich regions of the genome. Driven by ambient thermal fluctuations, a DNA bubble is a dynamical entity whose size varies by thermally activated zipping and unzipping of successive base pairs at the two forks where the ssDNA bubble is bordered by the dsDNA-helix. This incessant zipping and unzipping leads to a random walk in the bubble-size coordinate, and to a finite lifetime of DNA-bubbles under non-melting conditions, as eventually the bubble closes due to the energetic preference for the closed state. This so-called DNA breathing typically opens up a few bps. It has been demonstrated recently that by fluorescence correlation methods the fluctuations of DNA bubbles can be explored on the single molecule level, revealing a multistate kinetics that corresponds to the picture of successive zipping and unzipping of single base pairs. At room temperature, the characteristic closing time of an unbounded base pair was found to be in the range 10 to 100 μsec corresponding to an overall bubble life time in the range of a few msec. The multistate nature of the DNA-breathing was recently confirmed by a UV-light absorption study.

The presence of fluctuating DNA-bubbles is essential to the understanding of the binding of single-stranded DNA binding proteins (SSBs) that selectively bind to ssDNA and that play important roles in replication, recombination and repair of DNA. One of the key tasks of SSBs is to prevent the formation of secondary structure in ssDNA. From the thermodynamical point of view one would therefore expect SSBs to be of an effectively helix-destabilizing nature, and thus to lower T_m. However, it was found that neither the gp32 protein from the T4 phage nor *E.coli* SSBs do. An explanation to this apparent paradox was suggested to consist in a kinetic block, i.e., a kinetic regulation such that the rate constant for the binding of SSBs is smaller than the one for bubble closing. This hypothesis could recently be verified in extensive single molecule setups using mechanical overstretching of dsDNA by optical tweezers in the presence of T4 gene 32 protein [1,2]. Our study consisted in developing a biophysical model of this system, as detailed below; see references [3-7], and references cited therein.

B. Probabilistic modeling by the master equation (ME)

The method has been outlined in detail in the previous technical report "Regulation of DNA metabolism by DNA-binding proteins probed by single molecule spectroscopy". The main features are as follows.

DNA breathing in presence of SSBS is described by the probability distribution $P(m,n,t)$ to find at time t a bubble with m open bps and with n SSB units bound to the two arches of the bubble. The time evolution of $P(m,n,t)$ is given by a master equation

$$\frac{\partial}{\partial t}P(m,n,t) = W\,P(m,n,t) \qquad (2)$$

where the transfer matrix W incorporates rates $t^{\pm}$ for bubble increase / decrease as well as rates $r^{\pm}$ for SSB binding / unbinding:

$$\begin{aligned}\frac{\partial}{\partial t}P(m,n,t) &= t^{+}(m-1,n)P(m-1,n,t) + t^{-}(m+1,n)P(m+1,n,t) \\ &- \left\{t^{+}(m,n) + t^{-}(m,n)\right\}P(m,n,t) \\ &+ r^{+}(m,n-1)P(m,n-1,t) + r^{-}(m,n+1)P(m,n+1,t) \\ &- \left\{r^{+}(m,n) + r^{-}(m,n)\right\}P(m,n,t)\end{aligned} \tag{3}$$

The rates $t^{\pm}$ and $r^{\pm}$ have contributions from the entropy of single-stranded DNA bubbles and from the binding of SSBs. The final result for the bubble size rates, $t^{\pm}$, are

$$\begin{aligned}t^{+}(m,n) &= k\,u\left(\frac{m+1}{m+2}\right)^{c}, \qquad m \geq 1 \\ t^{-}(m,n) &= k\,\frac{\Omega(m-1,n)}{\Omega(m,n)}\end{aligned} \tag{4}$$

where $c = 1.76$ and k is the base pair zipping rate. The SSB number transfer rates, $r^{\pm}$, result as

$$\begin{aligned}r^{+}(m,n) &= \gamma\,k\,\kappa\,(n+1)\,\frac{\Omega(m,n+1)}{\Omega(m,n)} \\ r^{-}(m,n) &= n\,\gamma\,k\end{aligned} \tag{5}$$

with the dimensionless ratio $\gamma = q/k$ of the SSB unbinding rate q and the base pair zipping rate k.

The master equation (3) can be solved by means of an eigenmode expansion of the form

$$P(m,n,t) = \sum_{p} c_p\,Q_p(m,n)\,\exp\left(t/\tau_p\right) \tag{6}$$

in which the coefficients c_p of a given eigenmode p are determined via the initial conditions. The corresponding eigenvalue equation for the bubble size-SSB number eigenfuction Q_p determines the mode relaxation times τ_p and can be solved numerically.

Fast SSB-binding corresponds to the limit in which the dimensionless ratio $\gamma = q/k$ of the SSB unbinding rate q and the base pair zipping rate k is large, i.e., $\gamma \gg 1$. This limit allows one to average out the SSB-dynamics and to calculate an effective free energy, in which the bubble dynamics with the slow variable m runs off. The result for two different binding strengths κ is shown in **Figure 3**, along with the free energies corresponding to keeping n fixed. It is distinct that while for lower κ the presence of SSBs diminishes the slope of the effective free energy, for larger κ the slope becomes negative. This implies that in the first case the bubble opening is more likely, but still globally unfavorable. In the latter case, the presence of SSBs indeed leads to full denaturation. One observes distinct finite size effects due to $\lambda > 1$: only when the bubble reaches a minimal size $m \geq \lambda$, SSB-binding may occur, a second SSB is allowed to bind to the same arch only once $m \geq 2\lambda$, etc. This effect also produces the nucleation barrier for full denaturation in the lower plot of **Figure 3**.

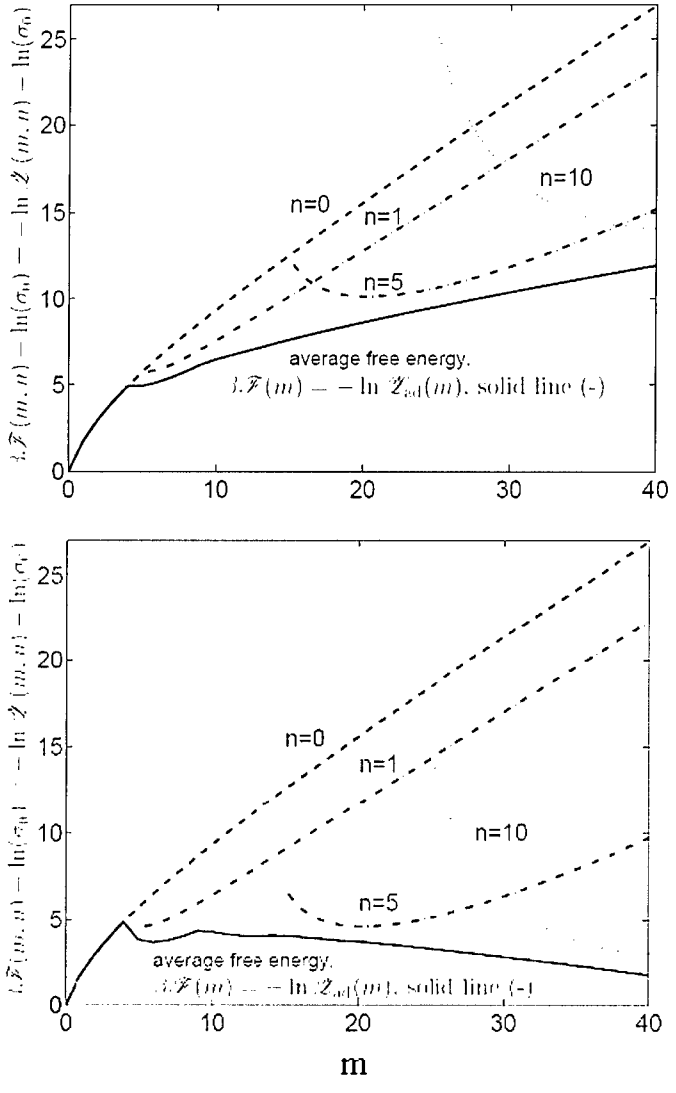

Figure 3. Effective free energy in the limit $\gamma \gg 1$ (solid line), and free energy for various fixed n ($u = 0.6$, $c = 1.76$, $\lambda = 5$). Top: $\kappa = 0.5$; bottom: $\kappa = 1.5$.

C. AFM-based evaluation of Low Density Lipoprotein interaction with DNA

In addition to the biophysical modeling outlined above we have establishing a procedure for AFM-based evaluation of Low Density Lipoprotein (LDL)-DNA interaction with DNA containing human cytomegalovirus (HCMV) promoter. This study is addressed in three parts: 1. imaging of DNA; 2. imaging of LDL; and 3. imaging of the LDL-DNA complex. Our effort was primarily directed towards elucidating optimal imaging conditions for imaging of DNA-protein complexes in air and in fluid.

We were successful in obtaining images of HCMV-containing DNA (plasmid pBR322, Promega) with bound proteins (**Figure 4**) in air. Although the identity of the proteins molecules observed on these images has not been established, we speculate that the particles aligned with DNA molecule are native E. Coli proteins (transcription factors) that remain associated with DNA during the purification of the plasmid from E-Coli cells.

Based on these images, we redesigned our approach to LDL-DNA binding studies: we will obtain plasmids with multiple LDL binding sites that would be easily identified by AFM.

We greatly appreciate professional advice from Dr. Steven Levene of University of Texas at Dallas.

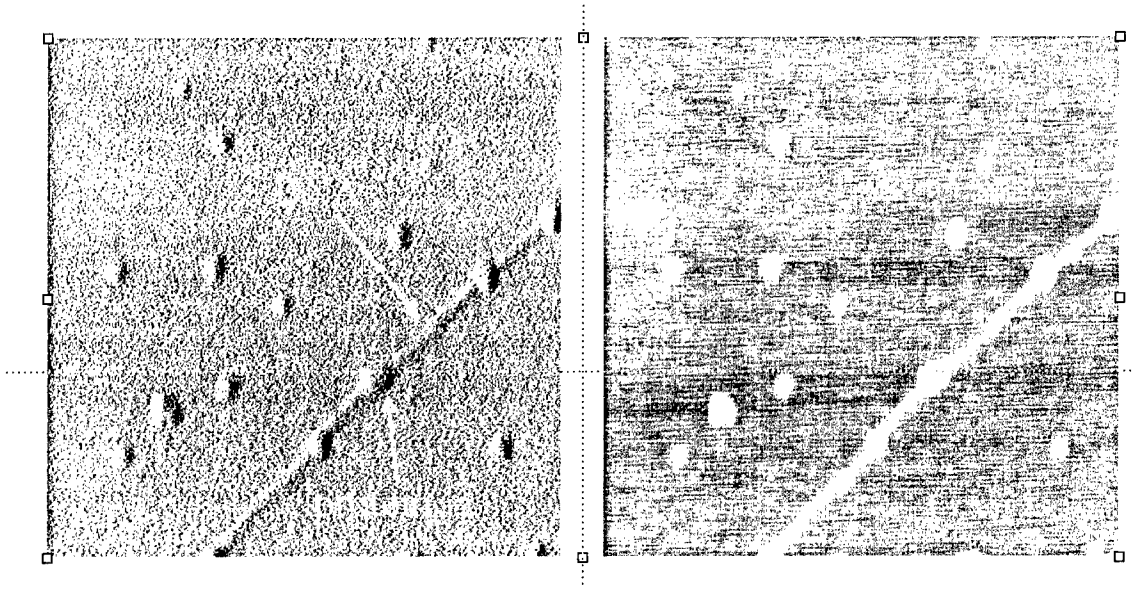

Figure 4. Plasmid DNA (pBR322) imaged using tapping mode AFM in air. Images are recorded by Mr. Juan Atkinson (Physics graduate student) using Multimode SPM with Nanoscope IV controller (Veeco Instruments) in the cantilever deflection amplitude mode (A) and the height mode (B).

D. References

[1] K. Pant, R. L. Karpel, I. Rouzina, and M. C. Williams, J. Mol. Biol. **336**, 851 (2004).

[2] K. Pant, R. L. Karpel, I. Rouzina, and M. C. Williams, J. Mol. Biol. **349**, 317 (2005).

[3] A. Hanke and R. Metzler, Phys. Rev. Lett. **90**, 159801 (2003).

[4] A. Hanke and R. Metzler, J. Phys. A **36**, L473 (2003).

[5] T. Ambjornsson and R. Metzler, J. Phys. Cond. Mat. **17**, S1841 (2005).

[6] T. Ambjornsson and R. Metzler, Phys. Rev. E **72**, 030901(R) (2005).

[7] R. Metzler, T. Ambjornsson, A. Hanke, Y. Zhang, and S. D. Levene, Journal of Computational and Theoretical Nanoscience (in press, 2006). E-print physics/0609139.