Modeling DNA unzipping in the presence of bound proteins

Farhat Habib and Ralf Bundschuh*

The Ohio State University, Department of Physics, 191 W Woodruff Ave, Columbus, OH 43210
(Dated: July 25, 2005)

Unzipping Force Analysis of Protein Association is a novel technique to investigate protein-DNA interactions by mechanically unzipping DNA. We computationally investigate the limits of this technique under quasi-static conditions. We find the minimum binding energy of a protein for which the protein can be detected using this technique and the minimum distance between the binding sites of two proteins of varying binding energies that can be resolved unambiguously with this technique.

I. INTRODUCTION

One of the most exciting developments in biophysics in recent years has been the advent of experimental techniques that allow mechanical experiments to be performed on single molecules [1]. Among the many techniques that have been developed for this purpose, such as Atomic Force Microscope cantilevers, glass microfibers, and optical or magnetic tweezers, optical tweezers are the best suited for studying protein-DNA interactions [2]. Optical tweezers have sensitivity in the picoNewton range and the distance at which the force is being applied can be measured to nanometer accuracy. The forces and distances encountered in DNA unzipping in the presence of bound proteins are within the same range as those offered by optical tweezers making it the ideal choice for performing mechanical denaturation experiments on protein-DNA systems [3].

A large number of cellular processes depend on DNA-protein interactions where the tasks that DNA participates in are mediated or catalyzed by DNA-binding proteins. These DNA-binding proteins such as polymerases, helicases, nucleases, isomerases, ligases, and histones play different and vital roles in gene expression, suppression, replication, transcription, recombination, repair and other cellular activities. To perform these tasks, proteins have to bind to the appropriate site on the DNA strand to facilitate its function [4, 5]. The binding sites of proteins can be highly sequence specific; thus, knowledge of binding sites can give us information on the sequence information in a DNA strand. While rapid DNA sequencing methods now allow us to know the sequence information easily, information about binding sites on the sequence is harder to find [6]. UFAPA infers the binding site of a protein on a DNA strand using the peaks in the force-extension curve obtained by unzipping the DNA strand [3]. This technique can also aid in the discovery of new DNA-binding proteins and their properties. In addition, the binding energy between the protein and DNA,

which can also be measured using this technique, is a valuable clue to the nature of protein-DNA interaction.

Previously, it has been known that the force required to unzip naked DNA depends strongly on the local nucleotide sequence [7]. We extend this to predict the amount of force required to unzip DNA with bound proteins of varying binding energies. Specifically, we quantify the minimum binding energy for a protein that can be observed using this single-molecule technique under quasi-static conditions. We then go on to find the minimum distance between two protein molecules that can be resolved unambiguously with this technique.

The remainder of this manuscript is organized as follows. We give a description of the model, the theory behind it, and a brief description of the methods used for the computations. Then we present the results obtained for the minimum binding energy that can be detected with UFAPA, and then finally we have the results for the minimum resolvable distance between two proteins of varying binding energies and a discussion of the results obtained.

II. MODEL

In the experimental setup that we want to model, the two strands of a DNA molecule are being pulled apart mechanically using a device with which we can measure the force and the distance between the two ends of the molecule being pulled apart as shown in Fig. 1. E.g., in the original UFAPA experiment [3], one end of the DNA molecule is held fixed using the gradient force from an optical tweezer, while the other end is attached to a glass coverslip that is gradually pulled away. The force $f(R)$ is measured as a function of the end-to-end distance $R = |\mathbf{R}|$, where $\mathbf{R}$ denotes the vector between the two ends of the DNA molecule. In practice, the force measurement requires a device that acts like a spring, hence the distance cannot be kept exactly constant but if we operate in a regime where the stiffness of this spring is much higher than that of the single-stranded (ss)DNA that has already been pulled out, we can ignore the presence of this spring [8].

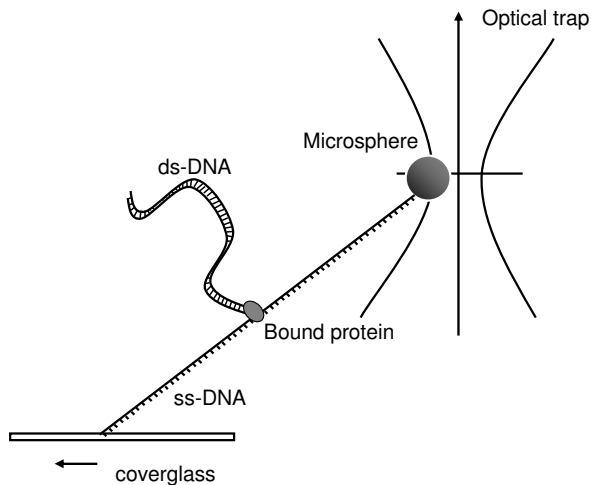To calculate the force required to keep the two ends at a

FIG. 1: A cartoon illustrating DNA being unzipped using optical tweezers.

fixed distance we use the partition function of the system. The partition function at fixed extension, $Z_N(R)$, for a given DNA sequence of N nucleotides, may be written as a sum over the number m of exterior open basepairs, i.e., the number of open basepairs until the first closed basepair as seen from the end of the molecule that is being pulled upon. For a fixed number m, we can split the partition function into two parts, the partition function, $\mathcal{Q}_N(m)$ from the $N - m$ basepairs starting from the first closed basepair and the partition function from the $2m$ open bases. The latter depends on the end-to-end vector $\mathbf{R}$ and is given by a suitably selected polymer model for the ssDNA. We will denote this partition function as $W(\mathbf{R}; m)$. Together, we obtain

$$Z_N(R) = \sum_m \mathcal{Q}_N(m) W(\mathbf{R}; m) \qquad (1)$$

Calculating $\mathcal{Q}_N(m)$: We assume that the closed part of the DNA sequence does not contain any bubbles, i.e., regions of open DNA between closed bases. This assumption is justified as the formation of bubbles is highly unfavorable around room temperature at which these single-molecule experiments are typically performed [9]. Only close to the significantly higher denaturation temperature of DNA at $60-90°$C bubbles start to become important. Then, the partition function $\mathcal{Q}_N(m)$ of the dsDNA, in the absence of any bubbles is given by

$$\mathcal{Q}_N(m) = e^{-\beta \sum_{i=m+1}^{N} E(i)} \qquad (2)$$

where $E(i)$ is the stacking energy of the basepair at position i. The stacking energies of a basepair depend on that pair and the neighboring pair resulting in 10 unique neighbor-pairs in DNA [10].

Polymer model: The simplest polymer model for the exterior ssDNA is a Gaussian chain [11]. However, at the force scales at which DNA unzipping occurs, the exterior strand is strongly stretched and the Gaussian model is of insufficient accuracy. Thus, we use an elastic freely jointed chain (EFJC) model which provides a better fit to experimental force-extensions curves (FECs) [12].

In the elastic freely jointed chain model the ssDNA is modeled by elastic segments of rest length l and a spring constant κ. The elastic energy per segment is then given by $V(\mathbf{r}) = \frac{\kappa}{2}(r - l)^2$, where $\mathbf{r}$ is the end-to-end vector of each segment. Instead of calculating the exact partition function $W(\mathbf{R}; m)$ of the chain, which is quite cumbersome, we employ an asymptotic expression that becomes exact in the limit of a large number of segments [13]. In practice the approximation used becomes very good already after a few tens of segments. Since in our application the single-stranded polymer consists of on the order of a thousand segments this approximation is sufficiently accurate for our purposes.

The individual bases in ssDNA are not completely free to move around as the EFJC model requires. To take this into account we use the Kuhn length of ssDNA which is the length scale at which the chain is effectively free. Thus, we model a strand of m opened basepairs by $m' = 2ml_b/l_K$ segments of the EFJC where the 2 comes from the fact that each bond opening results in 2 segments to the open chain. Here l_b is the base length and l_K is the Kuhn length of ssDNA. The asymptotic expression can then be written in terms of the function $q(h)$ as [13]

$$W(\mathbf{R}; m) \approx C \frac{h}{2\pi R} [q(h)]^{2ml_b/l_K} e^{-hR} \qquad (3)$$

where C is a normalization constant. The function $q(h)$ itself is given by

$$q(h) = \frac{\int d^3 r e^{-\mathbf{h} \cdot \mathbf{r} - V(\mathbf{r})/k_B T}}{\int d^3 r e^{-V(\mathbf{r})/k_B T}} \qquad (4)$$

where k_B is the Boltzmann constant and $\mathbf{h}$ is a vector with some arbitrary orientation in space and a length h that is determined from $R = m \frac{2l_b}{l_K} \frac{\partial}{\partial h} \log q(h)$. To simplify $q(h)$ we write it as $q(h) = I(h)/I(0)$, where $I(h) = \int d^3 r e^{-\mathbf{h} \cdot \mathbf{r} - \mathbf{V}(\mathbf{r})/\mathbf{k_B T}}$. Using spherical polar coordinates and integrating over the angular coordinates we reduce $I(h)$ to

$$I(h) = 2 \int dr \frac{r}{h} e^{-V(r)/k_B T} \sinh(hr) \qquad (5)$$

which can be solved in terms of hypergeometric functions.

Modeling the protein: The protein-DNA binding is modeled by including an extra term for the protein-DNA binding energy in the partition function. It depends on a single additional parameter, namely the effective free energy of binding ΔG_{eff}. This free energy of binding is given by

$$\Delta G_{\text{eff}} = k_B T (\ln[\text{protein}] + \ln K_a) \qquad (6)$$

where

$$K_a = \frac{[\text{protein.DNA}]}{[\text{protein}][\text{DNA}]} \qquad (7)$$

is the equilibrium association constant for a single binding site. Thus, the effective free energy of binding depends on the protein concentration as well as the chemical binding free energy of the protein-DNA complex.

With this, the partition function for a DNA sequence with a single protein molecule becomes

$$Z_N(R) = \sum_m \mathcal{Q}_N(m)W(\mathbf{R};m)(1+\theta(m-m_0)e^{\Delta G_{\text{eff}}/k_B T})$$

(8)

where θ is the Heaviside step function, and m_0 is the base position at which the protein binds to the DNA. We use $1+\theta(m-m_0)e^{\Delta G_{\text{eff}}/k_B T}$ in order to account for the possibility of protein being bound or unbound. This can easily be extended for two or more proteins bound to the DNA strand. In this paper we use kJ/mol as the units for the binding energy (1 kJ/mol = $0.406 k_B T$/molecule at 293 K).

Here we are assuming that the process of DNA unzipping is a quasi-static or reversible process, that is, the time scale over which protein binding/unbinding takes place is much faster compared to the unzipping of the DNA. Experimentally, this means that if we run the unzipping in reverse, that is "zipping" instead of unzipping the DNA, we should obtain the same curve. Presence of hysteresis would indicate that the experiment was not over a time scale long enough to be considered quasi-static. While DNA binding or unbinding is fast enough to be considered quasi-static for the original UFAPA experiment [3], the process of protein association or disassociation is considerably slower and a quasi-static approximation is a significant oversimplification. Thus, to obtain more accurate results we would have to consider the kinetics of protein unbinding. This would minimally require explicitly modeling the dynamics of the protein as a two-state model with on and off rates given by the microscopic on and off rates modified by the probability of the binding site being present as given by the thermodynamics of the unbinding DNA assumed to be in thermal equilibrium. Such a treatment would require the knowledge of the two microscopic rates which are usually much more elusive than the equilibrium constants used here. Thus, we consider our quasi-static modeling presented here as a first step toward modeling the UFAPA experiment.

Force-extension curves: Once we have the partition function $Z_N(R)$, the force at fixed extension R is given by:

$$f(R) = -k_B T \frac{\partial}{\partial R} \log Z_N(R)$$

(9)

This force $f(R)$ is the average force at fixed extension R. In a single-molecule experiment at fixed extension R, the actual force will fluctuate around this value but this value will emerge if these fluctuations are averaged over sufficiently long time intervals.

Implementation: Since the partition function at each point in the curve has to be summed over all basepairs N, we can see that this rapidly becomes unwieldy

for large sequences. To obtain the curves in a reasonable time we modify the calculation of the partition function as follows. The function $\mathcal{Q}_N(m)W(\mathbf{R};m)$ is sharply peaked around $m = R/l_b$, where l_b is the base length of the ssDNA. Thus, while summing over all m only ms close to the peak have a significant contribution as can be seen from Fig. 2.
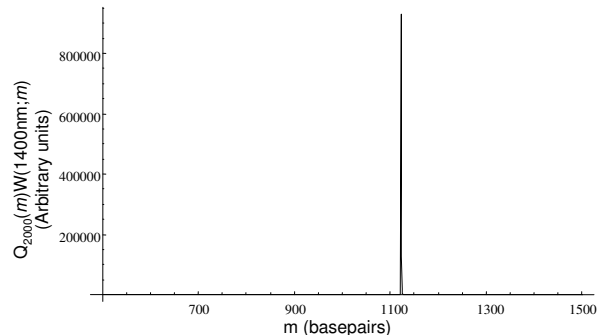

FIG. 2: A plot showing the (unnormalized) contribution of each index m toward the partition function at a fixed extension $R = 1400$ nm and a DNA length of 2000 basepairs.

Therefore, we start the summation from the peak thus identified and stop when the value of $\mathcal{Q}_N(m)W(\mathbf{R};m)$ falls to less than 10^{-9} times the value at the peak on both sides. This significantly reduces the number of computations that need to be performed. The relative difference between the force-extension curves from the restricted partition function and the exact partition function was found to be less than 10^{-11} for the extension considered for bare DNA and less than 0.005 in the presence of proteins. All calculations were performed using Mathematica [14].

Parameters: The elastic parameters that we used for ssDNA were [12]: a contour length per base (l_b) of 0.539 nm, a persistence length ($l_p = l_K/2$, where l_K is the Kuhn length) of 0.796 nm, and a stretch modulus of 580 pN. The stacking energies for DNA were obtained from Allawi and SantaLucia Jr. [15]. We used random DNA sequences with varying proportion of ATs and GCs under ionic conditions of 0.1M Na$^+$ to have the same conditions as Koch et al. [3].

III. RESULTS AND DISCUSSION

There are several interesting and experimentally relevant questions about this technique that can be resolved computationally. The questions that this manuscript answers are what is the minimum energy for the protein-DNA complex for which we can still reliably detect the bound protein and what is the minimal distance in basepairs over which two bound proteins can be separately resolved.

Before we address these questions we verify that our model reproduces results that have been observed experimentally for DNA in the absence of proteins. For such bare DNA having an equal proportion of ATs and GCs we find an average unzipping force of 15.3 pN with a standard deviation of 0.7 pN. This corresponds well with experimental measurements of 12-17 pN for bare DNA under the chosen ionic conditions [3]. The force extension curves obtained for one specific randomly chosen DNA sequence can be seen in Fig. 3 along with the mean force and the standard deviation. The values obtained for the mean force and the standard deviation are weakly dependent on the end-to-end distance of the DNA molecule. All simulations performed here are for extensions R ranging from 1000 to 2000 nm. As the end-to-end distance R between the unzipped single strands increases, and their compliance thus decreases we find that there is a small decrease in the average unzipping force as well as the force required for protein-DNA dissociation. However, this decrease was found to be too small to be experimentally relevant hence we do not deem it necessary to investigate the dependence of our results on the extension in any detail.
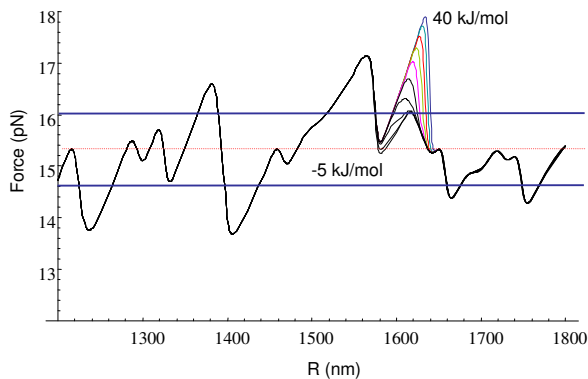

FIG. 3: (Color online) A series of calculated force-extension curves obtained with a DNA sequence with proteins of progressively less binding energy attached to the DNA molecule at the same position. The dotted line shows the mean unzipping force for a DNA strand in the absence of proteins. The solid lines indicate one standard deviation above and below the mean. The energies labeled are the binding energies of the protein for the highest and lowest curves. The lowest curve coincides with the curve for bare DNA.

A. Minimum binding energy

Once we add proteins to the DNA sequence the force extension curve changes as the unzipping fork reaches the binding site. The protein bound at the binding site prevents further unzipping from taking place resulting in a dramatic increase in force as the DNA sequence is stretched. Then as the protein unbinds there is a sudden relaxation and the unzipping force returns back to the bare DNA level. The force at which the protein dissociates from the DNA is directly related to the binding energy between the protein and the DNA. It also depends on the composition of the DNA sequence itself. Hence, we would like to know the relation between the height of a peak and the binding energy of the protein-DNA complex which can tell us at what binding energy the peak from a protein would be lost in the noise in the measurements. Effective protein binding energies generally vary between 5 kJ/mol to 120 kJ/mol depending on the protein, the DNA sequence, and the protein concentration. E.g., on the high side, the binding energy of cAMP receptor protein with lacDNA is 118.3 kJ/mol under usual conditions, while on the lower side, the binding energy of TATA binding protein with 20mer DNA is only about 6.7 kJ/mol [16].

Fig. 3 shows force-extension curves for a protein of progressively lower binding energy attached at the same position on a random DNA sequence. We start with a binding energy of 40 kJ/mol and decrease it in steps of 5 kJ/mol until we reach −5 kJ/mol. We notice that once the difference between the peak in the force extension curve due to the protein and the bare DNA is less than one standard deviation (SD) in the bare DNA itself then it is no longer possible to distinguish whether a peak came from a protein or is due to natural variation in the DNA itself. This happens at binding energies around 10 kJ/mol. As can be seen from the force-extension curve though, the peaks from the DNA itself are often above one SD from the mean, so we can see that at 2 SDs from the mean, it is already becoming very hard to distinguish between peaks from the protein and from the DNA sequence. This happens at binding energies of roughly around 20 kJ/mol. Thus, UFAPA will not be able to detect bound proteins below this binding energy under quasi-static conditions.

However, if we have knowledge of the force-extension curve from the bare DNA, then we can do better than 20 kJ/mol. We can take a difference between the curve with protein and the curve without protein to locate the peaks due to a bound protein. Given that the exact height of the peak depends not only on the protein binding energy but on the DNA sequence itself we averaged the force needed for protein dissociation over 15 random sequences each for GC:AT content in the sequence being 1:3, 1:1, and 3:1. Fig. 4 shows a plot of average force needed for protein dissociation above the force required for unzipping a bare DNA strand. With a knowledge of the force-extension curve of the bare DNA, the minimum detectable binding energy is essentially limited by the experimental noise in the system. For example, if the experimental error in the system is of the order of 0.5 pN, the minimum binding energy that could be detected would be 5 kJ/mol.
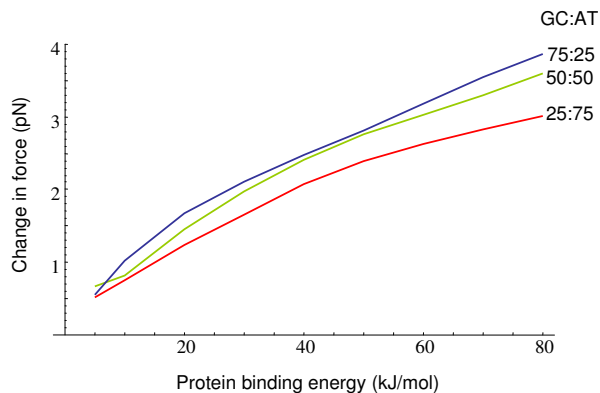
FIG. 4: (Color online) The average variation in the height of the force-extension curve peak above the unzipping force for bare DNA for DNA strands with varying GC:AT ratios. The heights were averaged over 15 random sequences each for a given GC:AT ratio

B.   Resolvable distance

Another proposed utility of UFAPA is the identification of protein binding sites on a given DNA molecule. Thus it is important to know the limits on the resolution of this technique when multiple proteins are present on the DNA sequence. We consider the force-extension curves obtained in the presence of two proteins on the DNA strand.
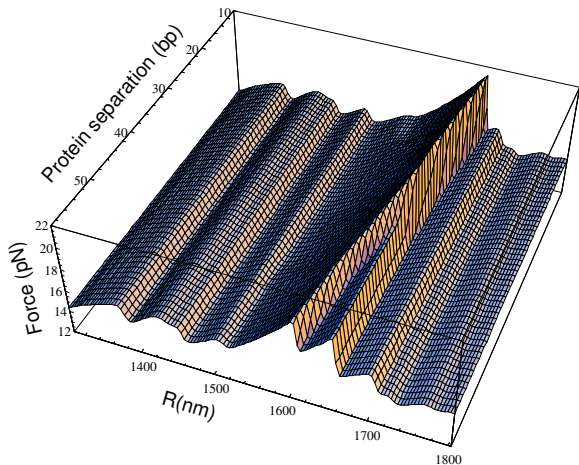

FIG. 5: (Color online) The surface formed by force-extension curves of a DNA molecule having two proteins of equal binding energy of 100 kJ/mol attached to it as the protein binding positions are separated further apart. Only a single peak is seen until the proteins are separated by a distance of roughly 35 basepairs. Thus, UFAPA will not be able to resolve two proteins closer than this distance.

Case study: The surface in Fig. 5 shows the variation of the force extension curves obtained when two proteins of equal strength are brought closer together. We see that when the proteins are quite distant from each other, there

are two distinct peaks as the proteins break off from the chain. However, as they come closer to each other the peaks come closer faster than the proteins themselves thus merging well before the proteins are at the same position. Thus, the second protein which does not seem to have any role to play at the unzipping fork is able to influence the amount of force it would take to dissociate the first protein.
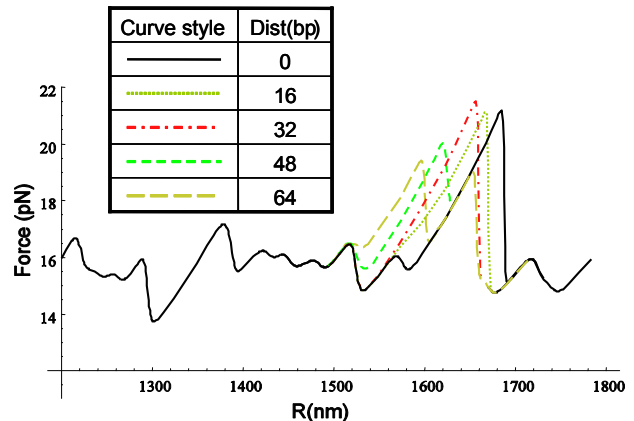

FIG. 6: (Color online) Force-extension curves from a DNA strand having two bound proteins initially separated by 64 basepairs. The series of curves are obtained by bringing the proteins closer in steps of 16 basepairs until they are at the same position. As can be seen, the peaks from the two proteins merge while the proteins are still separated by more than 30 basepairs.

We want to stress that we do not take into account any direct interaction between the proteins in our calculation. The apparent interaction between the two proteins is exclusively mediated through the thermodynamics of the unzipping DNA. Any attractive direct interaction (which would depend on details like the geometry of the proteins bound at a certain distance from each other) would enhance the cooperativity observed here.

In order to allow for a more quantitative investigation, a selection of some of the FECs from Fig. 5 are shown together in Fig. 6. There are two striking features of the two protein-DNA system. First, the distance and force at which the second protein dissociates stays constant until the peaks from the two proteins have merged. Once merged, there is no difference between the curves obtained by two proteins of lower strength or a single protein of appropriate higher strength placed at the position of the first protein. The strength of this single protein is not a simple function of the strengths of the two proteins and the distance between them. Second, the maximum breaking force is obtained when the peaks from the two proteins have just merged into one. As the two proteins come closer together the breaking force decreases. This phenomenon can also be seen in a simpler model where the ssDNA is modeled by a Gaussian chain and we assume a uniform DNA sequence.

While at first glance this effect seems strange, it can be

understood by considering the effect the second protein has on the partition function. The partition function is a measure of the number of states accessible to the system [17]. The second protein restricts the number of states that the system can access for a given energy. The partition function $Z_N(R)$ is composed of 3 parts, the part due to the dsDNA $\mathcal{Q}_N(m)$ which is dominant at small R, the part due to the ssDNA $W(\mathbf{R}; m)$ which is dominant at large R, and the term due to the protein which is dominant when R is such that the DNA is bound up to the protein but not beyond that, i.e. the bound protein is restricting the dsDNA from further unzipping. Before the protein is encountered by the unzipping fork, the log of the partition function decreases roughly linearly with distance resulting in an almost constant slope of $\log Z_N(R)$ with respect to R. With the addition of a protein, the extra energy term from the protein in the partition function causes the partition function to decrease faster than the linear decrease from the distance where the protein is encountered to some distance beyond until the additional term from the protein is overwhelmed by the increasingly dominant $W(\mathbf{R}; m)$ term. On the FEC, given by $-k_B T \frac{\partial}{\partial R} \log Z_N(R)$, this appears as a rise in the curve followed by a sudden drop as the protein dissociates. When a second protein follows the first protein closely, the already decreased slope of the partition function meets a further decrease due to the second protein before the slope can return to its original value by the unzipping of the first protein. Due to this, the peaks due to the unzipping of the two proteins merge and appear as one.
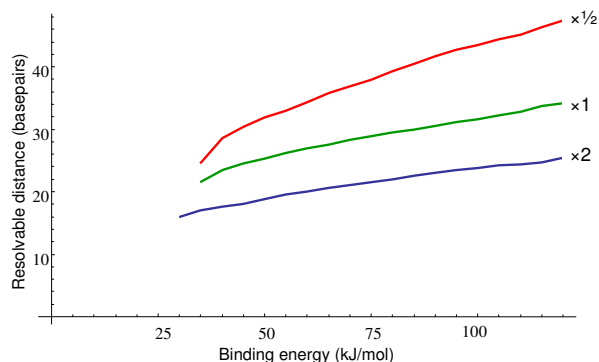

FIG. 7: (Color online) The minimum resolvable distance between two proteins on a DNA strand using UFAPA. The number after each curve gives the ratio between the second protein to be unbound and the first protein to be unbound. The x-axis indicates the strength of the first protein. The resolvable distances are averaged over 15 random sequences.

Generic trends: Within our computational model, we can quantify how the resolvable distance depends on the binding energies of the two proteins. However, identifying the minimum resolvable distance for different protein binding energies over a number of sequences manually is tedious and time-consuming. To automate this process, we use the first feature noted earlier, that the force and distance at which the second protein dissociates does not change until the individual peaks due to the two proteins in the force extension curve have merged into one. This suggests the algorithm that we use to calculate the resolvable distance for a range of protein strengths over many DNA sequences.

We start at some distance R that is beyond the point where the second protein is expected to dissociate. We gradually reduce R until there is a sharp rise in the force-extension curve which tells us that we have reached the distance at which the second protein dissociates. We record the force at this point and move the first protein closer by one basepair at a time. When the peaks merge there is a change in the force at this distance R. At this point we record the distance between the positions of the two proteins. We repeat this procedure for different protein strengths and DNA sequences as needed.

Fig. 7 shows the minimum distance at which two bound proteins can be resolved versus the protein strengths. The x-axis represents the binding energy of the first protein to be unzipped. The number after the curve gives the ratio between the binding energies of the second protein and the first protein. These curves have been averaged over 15 random sequences to remove the effect of the underlying DNA sequence on the resolvable distance.

We notice that as the effective binding energy of the first protein increases with respect to the second protein there is an increase in the distance at which we can resolve the peaks from the two proteins separately. Two proteins closer than this distance, will appear as a single protein with a higher binding energy on the force-extension curves from UFAPA in the quasi-static regime.

IV. CONCLUSIONS

We computationally investigated the limits on the detection of protein-DNA interactions using UFAPA. We found that in the quasi-static limit the minimum effective protein binding energy for a single protein that can be detected is roughly 20 kJ/mol in the absence of a FEC from the bare DNA. We found that the height a peak due to presence of a bound protein varies from 1 pN to 4 pN above the force required for unzipping the bare DNA strand, as the protein binding energy varied from 10 kJ/mol to 80 kJ/mol. If the FEC of the bare DNA is known, the minimum detectable binding energy is determined by the experimental error in the system. There is a dependence on the GC content of the DNA sequence as well. We also found the minimum distance between two proteins for which they can be resolved unambiguously. We find that resolvable distances between two proteins ranges from ~ 10 basepairs to as much as 50 basepairs depending on the binding energies of the two proteins. As compared to the experiment, we find that the force at which a protein dissociates is higher than the force found in our simulations. This is because the unzipping in the experiment is too fast to be considered

quasi-static. Thus, it will be crucial to consider the kinetics of the protein unbinding process as the next step of modeling and understanding this important experimental technique.

[1] A. D. Mehta, M. Rief, J. A. Spudich, D. A. Smith, and R. M. Simmon, Science **283**, 1689 (1999).
[2] R. H. Austin, J. P. Brody, E. C. Cox, T. Duke, and W. Volkmuth, Physics Today **54**, 32 (1997).
[3] S. J. Koch, A. Shundrovsky, B. C. Jantzen, and M. D. Wang, Biophysical Journal **83**, 1098 (2002).
[4] G. M. Cooper, *The cell, A Molecular Approach* (ASM Press, The American Society of Microbiology, Washington, DC, 1997).
[5] A. Kornberg and T. A. Baker, *DNA Replication. 2nd ed.* (W. H. Freeman, New York, 1992).
[6] G. D. Stormo, Bioinformatics **16**, 16 (2000).
[7] U. Bockelmann, B. Essevaz-Roulet, and F. Heslot, Physical Review Letters **79**, 4489 (1997).
[8] U. Gerland, R. Bundschuh, and T. Hwa, Biophysical Journal **81**, 1324 (2001).
[9] Y. Kafri, D. Mukamel, and L. Peliti, The European Physical Journal B **27**, 135 (2002).
[10] R. L. Ornstein, R. Rein, D. L. Breen, and R. D. MacElroy, Biopolymers **17**, 2341 (1978).
[11] P.-G. de Gennes, *Scaling concepts in polymer physics* (Cornell University Press, Ithaca, NY, 1979).
[12] S. B. Smith, Y. Cui, and C. Bustamante, Science **271**, 795 (1996).
[13] P. Flory, *Statistical Mechanics of Chain Molecules* (Interscience Publishers, New York, 1967).
[14] W. R. Inc., *Mathematica 5.0* (Wolfram Research, Inc., Champaign, IL, 2003).
[15] H. T. Allawi and J. SantaLucia Jr., Biochemistry **36**, 10581 (1997).
[16] P. Prabakaran, J. An, M. M. Gromiha, S. Selvaraj, H. Uedaira, H. Kono, and A. Sarai, Bioinformatics **17**, 1027 (2001).
[17] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Addison-Wesley, 1969).